Big-picture ecologist generates buzz and backlash p. 412

Race-blind algorithms cause racial disparities pp. 421 & 447

Tomorrow's Earth pp. 422, 443, & 444

Science

$15
25 OCTOBER 2019
sciencemag.org

AAAS

WORMS OF THE WORLD

The global diversity and biogeography of earthworms pp. 425 & 480

CONTENTS

25 OCTOBER 2019 • VOLUME 366 • ISSUE 6464

409 & 430

The raccoon-size *Loxolophus* and other mammals evolved surprisingly quickly after the end-Cretaceous extinction.

NEWS

IN BRIEF

IN DEPTH

FEATURES

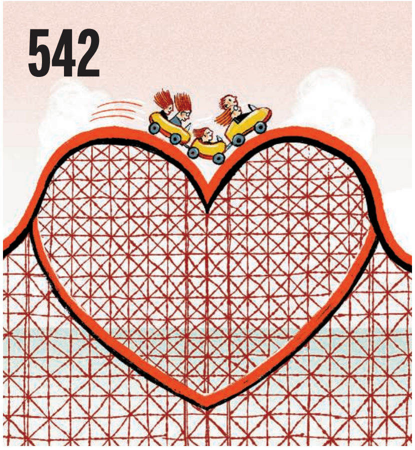
INSIGHTS

PERSPECTIVES

POLICY FORUM

ILLUSTRATION: HHMI TANGLED BANK STUDIOS

Published by AAAS



ON THE COVER

Earthworms are major components of our world's soils and drive important ecosystem services. However, the main factors affecting their global diversity and abundance have been largely unknown until now. Climate is the main environmental factor driving earthworm biodiversity patterns, with higher local diversity and abundance occurring in temperate regions. Despite the lower local diversity in the tropics, across the entire region there may be more species overall. See pages 425 and 480.
Photo: Picture Partners/Alamy Stock Photo

ILLUSTRATION: ROBERT NEUBECKER

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2019 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1971; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery): $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

EDITORIAL

On privilege

This editorial is my last contribution in my role as Editor-in-Chief of the *Science* family of journals. I have had a front-row seat to learning about often astounding research across many disciplines. And I have overseen efforts at three mature journals and helped develop or launch three new ones. As I take stock of this tremendous opportunity, I have been reflecting on the characteristics of privilege, in science and in life.

I grew up on a college campus where my father was a mathematics professor and my mother was a physician involved in research. This provided me with encouragement for academic pursuits and access to many resources. I greatly appreciated this atmosphere, but it was when I started college that I remember first articulating an important aspect of my privilege. I realized that my fellow students were intimidated by faculty members and said to my roommate, "It's hard to be too afraid when you've seen the famous professor down the street shoveling dog poop off his lawn." This seemingly trivial privilege had huge consequences. Early in my freshman year, I reached out to an assistant professor about potential opportunities in his laboratory. A month later, he invited me to help with a collaborative project, initially with menial tasks—tracing computer output onto plastic sheets—but it led to opportunities to learn about the project and to become assimilated into the laboratory environment. Over a short period of time, I was helping with other aspects of the research and handling my own project. More opportunities followed, as did authorship on research publications. My experience illustrates a central aspect of privilege—positive feedback loops. An initial advantage can lead to opportunities that, if capitalized on, can lead to additional ones.

"I have had a front-row seat to learning about often astounding research across many disciplines."

As a white male, I rarely had to deal with racism or sexism personally. I also did not have to face the antisemitism that my parents, particularly my mother, had to confront. My family was relatively secure financially. I did need to earn money to help with my education, but I never felt exposed—I had a safety net. Although my childhood had many positive aspects, it was not idyllic. Both my parents had mental health problems that often dominated their lives. I lived with the disruptions caused by these issues. I was fortunate to have a brother, 15 months my senior, who shouldered much of the burden of dealing with my parents' needs, while I coped by spending time at neighbors' houses and playing organized sports. Because of the privilege of this support, these adverse events helped me become self-reliant, learn on my own, and develop problem-solving skills.

Over the past 3 years, I have worked with the editors of the *Science* journals to break positive feedback loops that can narrow opportunities more than is optimal for science and society. We have strived to increase the diversity of our advisory boards in all dimensions. We have collected and analyzed data regarding the diversity of authors and reviewers and have acquired tools that can help track editorial decisions over time. By spreading opportunities across a wider and more diverse set of individuals with different life experiences and interests, we hope to strengthen the scientific enterprise now and moving into the future.

This is very much a work in progress, and I am delighted that Holden Thorp, a colleague of many years, will bring his wide range of experiences and leadership to the Editor-in-Chief position. He joins a group of editors who take the responsibility of fairness and equity very seriously and are eager to create a more equitable future for science. Holden brings his own privileges to the position but also a longtime commitment to promoting diversity that equips him well to lead this team.

–Jeremy Berg

Jeremy Berg
Editor-in-Chief,
Science Journals.
jberg@aaas.org

PHOTO: TERRY CLARK

10.1126/science.aaz8977

NEWS

IN BRIEF

Edited by **Jeffrey Brainard**

ARCHAEOLOGY

Sarcophagus cache found in Egypt

Egyptian scientists said last week that they uncovered 30 ancient sarcophagi near the Al-Asasif cemetery in Luxor, Egypt, reportedly the largest cache of coffins discovered in the country since the 19th century. The intricately painted coffins (above) date to about 1000 B.C.E. and contain the well-preserved mummies of men, women, and children. The sarcophagi were stacked in two layers and buried in sand rather than tucked into tombs—an unusual funerary ritual that may have kept them safe from looters. Future research on the mummies could reveal new information about the identities and health of elite Egyptians during the Third Intermediate Period, a time of political fragmentation after an age of expansion during the New Kingdom.

Permafrost exports carbon

CLIMATE SCIENCE | As it thaws from the effects of global warming, Arctic permafrost is releasing more carbon than it absorbs from new plant growth—adding to the atmosphere's burden and accelerating climate change, researchers report this week. Scientists have warned that the frozen soil represents a ticking bomb, but until now, Arctic studies have been too limited to predict when that shift would come. After pooling their observations from more than 100 Arctic field sites, scientists from the Permafrost Carbon Network estimated that permafrost released an average of 1.7 billion tons of carbon each winter from 2003 to 2017—double past estimates, they write in *Nature Climate Change*. Meanwhile, during the summer growing season, other surveys have found that the landscape absorbs only 1 billion tons—leaving an average of more than 600 million tons of carbon to escape to the atmosphere each year. The study remains limited by the paucity of Arctic observations; the overall uncertainty of Arctic winter emissions, for example, is plus or minus 813 million tons, nearly half the total emissions. The study also found no rise in emissions since 2003.

Reversal on Alzheimer's drug

CLINICAL RESEARCH | Shocking investors and scientists, drugmaker Biogen said it will ask the U.S. Food and Drug Administration to approve aducanumab, an Alzheimer's drug candidate that the company said had failed in March. On 22 October, Biogen said a new analysis of two clinical trials that studied patients with early-stage Alzheimer's disease showed that the drug significantly slowed clinical decline, primarily in those receiving the higher of two doses. The company said its March decision to abruptly end the trials was based on an interim analysis of results from 1748 trial subjects who took the drug for 18 months, which suggested it was unlikely to benefit patients; the new analysis included 2066 such patients. Aducanumab targets brain deposits of the protein beta-amyloid; several other drugs targeting beta-amyloid have been dropped by their developers.

House panel backs integrity bill

POLITICS | A 3-year effort to enact stronger protections for federal scientists against political interference took a big step forward last week, after Democratic sponsors of the bill struck a deal with Republicans. The science committee of the U.S. House of Representatives voted 25 to six to advance the Scientific Integrity Act, which would require some two dozen federal research

PHOTO: MOHAMED ABD EL GHANY/REUTERS

agencies to develop and follow such rules. Republicans saw the measure as an attack on the policies of President Donald Trump, whereas Democrats insisted that they were simply trying to ensure that federal scientists could pursue important research, publish the results, and discuss their findings at scientific conferences and with the public. The key compromise dropped language requiring agencies to allow scientists to talk freely to the media and instead gave individual agencies leeway to set guidelines. A companion bill is languishing in the Senate.

Space radiation mission ends

SPACE WEATHER | A NASA mission that explored the Van Allen belts—rings of charged particles trapped by Earth's magnetic field—ended its 7-year run last week. The second of the two Van Allen Probes ran out of fuel and was turned off, after its sister instrument ceased operations in July. The mission helped scientists discover many properties of the inner and outer belts, which protect Earth's atmosphere by capturing damaging, high-speed particles from the Sun and cosmic rays; the findings include a third belt that occasionally appears during intense solar storms. The probes, built to withstand radiation in a region of space that astronauts and other satellites avoid, spent more time in the belts than any other mission. NASA is preparing a successor mission, the Geostationary Transfer Orbit Satellite, expected to launch in 2021.

NASA's microbe rules faulted

PLANETARY SCIENCE | NASA rules that govern the potential spread of earthly microbes to other planets—and the potential return of alien life back to Earth—are anachronistic and should be relaxed, says a report released last week by an independent agency advisory panel. Although such planetary protection efforts remain a worthy goal, protecting other planets from all contamination is costly and does not make sense given current scientific knowledge, the panel said. The report suggests that NASA identify less vulnerable regions of the Moon and Mars, opening them up to increased exploration from humans and robots, and adopt modern genomic tools to characterize the microbes inhabiting its spacecraft and cleanrooms. Similarly, NASA should revisit the biohazard restrictions it places on samples returned from Mars to Earth, the report says, because martian meteorites have long bombarded Earth with no known biological harm.

IN FOCUS Most geodes—hollow, crystal-lined rocks—can fit in the palm of your hand. But one of the world's biggest, the Pulpí Geode (above)—discovered 20 years ago in an abandoned mine near its namesake town in southeastern Spain—measures 11 cubic meters. Last week in *Geology*, researchers described how the geode grew. They credit its great size in part to a plentiful supply of dissolved sulfate minerals from surrounding rocks and a lengthy time to grow.

PHOTO: HECTOR GARRIDO

Quake alerts for all Californians

SEISMOLOGY | The ShakeAlert earthquake early warning system, previously available only for residents of Los Angeles, will now deliver alerts to all Californians, the U.S. Geological Survey and its partners announced last week. The system broadcasts cellphone notices similar to weather and AMBER alerts to warn people of shaking from damaging earthquakes greater than magnitude 5. The MyShake phone app will provide more frequent alerts for those who want warnings of earthquakes of magnitude 4.5 and greater. Although ShakeAlert can provide tens of seconds of warning for some earthquakes, it cannot report a tremor's precise strength until seconds before it arrives (*Science*, 2 November 2018, p. 514). The warning system will eventually expand to two other seismically active states, Oregon and Washington.

IN OTHER NEWS

SCIENCE ADVISERS Thirty-three months after taking office, President Donald Trump has announced the first seven members of his President's Council of Advisors on Science and Technology. Only one is an academic; five hold Ph.D.s. Details at https://scim.ag/WHadvisers.

EBOLA VACCINE The world's first Ebola vaccine took a critical step closer to the market as a committee of the European Medicines Agency in Amsterdam recommended conditional approval.

CHILE TURMOIL Scientists in Santiago scrambled to maintain lab experiments after riots erupted there and in other cities over metro fares. Chile's Atacama Large Millimeter/submillimeter Array, the world's largest radio telescope facility, temporarily reduced operations.

Military personnel cover a potentially contaminated police car after the 2018 poisoning of Sergei Skripal and his daughter Yulia in the United Kingdom.

PHOTO: ANDREW MATTHEWS/AP PHOTO

CHEMICAL WEAPONS

Obscure Cold War nerve agents set to be banned

"Novichoks," used in 2018 attack on former spy, to come under chemical weapons treaty

By **Richard Stone**

The poisons were so fearsome that U.S. government scientists were forbidden from publicly uttering their name. Then, in 2018, one of the Novichok compounds was used in an attempt to assassinate a former Russian spy on U.K. soil—spurring the United States and allies to lift the veil of secrecy and mount a drive to outlaw the obscure class of nerve agents, concocted in a Soviet weapons lab during the height of the Cold War. Now, their effort to amend the Chemical Weapons Convention (CWC) is about to pay off.

On 9 October, the Executive Council of the Organisation for the Prohibition of Chemical Weapons (OPCW), the body that administers the treaty, reviewed a revised proposal from Russia that would bring Novichoks under the treaty's verification regime, along with a class of potential weapons known as carbamates. If the Russian proposal and a similar one from the United States, Canada, and the Netherlands are approved at a treaty review meeting next month, as expected, they would be the first update to the list of banned chemical weapons since the CWC came into force in 1997. "This is a historic milestone for the treaty," says Gregory Koblentz, a chemical and biological weapons expert at George Mason University in Fairfax, Virginia.

The newfound glasnost on Novichoks, also known as fourth-generation nerve agents, should spur research on their mechanism of action and on countermeasures and treatments. "Fourth-generation agents are now on the list of compounds we can study," says David Jett, director of the Countermeasures Against Chemical Threats Program at the U.S. National Institutes of Health. The U.S. government limits work on Novichoks to a handful of defense labs, but academic researchers may now partner with these labs and conduct computer modeling or other studies that don't require the chemicals. Such research, Jett hopes, will "provide more information on the toxicity of these threat agents."

Like other nerve agents, the known Novichok agents bind to acetylcholinesterase, an enzyme that dismantles the neurotransmitter acetylcholine after it's released into synapses. Without rapid medical intervention, the buildup of acetylcholine blocks brain signals from reaching muscles that control respiration and maintain blood pressure. Distinctive chemical groups jutting from the Novichok molecules may allow them to bind to other enzymes—and perhaps trigger a long-lasting syndrome in victims who survive an attack.

Chemical weapons experts had been whispering about Novichoks for decades. The first public clues came from Vil Mirzayanov, a Soviet military chemist who divulged the Novichok program in 1992. In a 2008 memoir, he revealed details about the chemicals' structures and claimed that some Novichok agents are several times more toxic than VX, the deadliest known nerve agent developed for warfare, which North Korean operatives used to assassinate Kim Jong-un's half-brother at the Kuala Lumpur airport in 2017. However, Mirzayanov did not publish toxicity data on A-234, the compound believed to have been used in last year's U.K. attack.

Treaty nations have long resisted adding Novichoks to the CWC's so-called Schedule 1 list of chemical weapons, which compels signatories to declare and destroy any stockpiles. "People were worried about a Pandora's box," fearing such a listing would force them to regulate ingredients of the weapons, Koblentz says. That could hamper the chemical industry and might clue in enemies on how to cook them up. (Who has the agents now is anyone's guess.) Indeed, the U.S. government for years classified the Novichok agents as top secret. "There was a desire among Western countries to keep the information as limited as possible to avoid proliferation issues," Koblentz says.

Last year's assassination attempt against former Russian spy Sergei Skripal in Salisbury, U.K., thrust the Novichok agents into the spotlight. The botched attack gravely sickened Skripal, his daughter Yulia, two police officers who investigated the crime scene, and a couple—Charlie Rowley and Dawn Sturgess—who a few months later happened on a perfume bottle containing the agent. After long hospitalizations, the Skripals, the officers, and Rowley recovered; Sturgess died. The United Kingdom charged two Russian men, reportedly military intelligence officers, as the alleged assail-

Published by AAAS

ants, and obtained a European warrant for their arrest; they remain at large in Russia.

Now, Novichok agents are shaping up as a potential area of comity. The Canadian-Dutch-U.S. proposal targets two broad groups of Novichoks, one of which includes A-234. The Russian proposal covers the same compounds and adds a third group of Novichoks and two families of carbamates. The United States studied carbamates, which also inhibit acetlycholinesterase, as potential weapons during the Cold War; some carbamates are reportedly twice as toxic as VX, note Koblentz and Stefano Costanzi, a nonproliferation expert at American University in Washington, D.C., in a review in the current issue of *The Nonproliferation Review*. (There are many less toxic carbamates as well, some of them in use as insecticides or drugs.)

Russia initially opposed the three-nation proposal to bring the Novichoks under the CWC, arguing it was scientifically "substandard" and politically motivated. But after months of wrangling, the sides have resolved their differences. In a sign of the new openness, the U.S. Department of Commerce in August published a description of both the Novichok agents and the carbamates that the two proposals would cover in the *Federal Register*—including structural information.

If nations approve the Schedule 1 listings next month, CWC negotiators could then choose to puzzle out which precursor ingredients to cite on other treaty schedules. "It's complicated," Costanzi says. "There could be many precursors, and it will be difficult to find an optimal solution that minimizes the risk of chemical proliferation without hindering legitimate industrial uses or revealing synthetic pathways for the preparation of Novichoks." But that task will be well worth the trouble. "It's not like Novichoks are relics of the past," Koblentz says. "These are weapons that are still killing people."

Other agenda items at next month's treaty meeting are far more contentious. Russia and Syria are expected to denounce the work of an OPCW panel tracing the origins of chemical weapons used in Syria's civil war and identifying the parties that used them. And Russia and its allies are likely to continue to resist efforts to eliminate a treaty provision permitting use of aerosolized incapacitants for law enforcement. Russia exploited that loophole in 2002, when it pumped two powerful opioids, remifentanil and carfentanil, into a theater in Moscow to subdue armed terrorists—killing scores of hostages along with the terrorists. ■

Richard Stone is senior science editor at the Howard Hughes Medical Institute's Tangled Bank Studios in Chevy Chase, Maryland.

GENETICS

Screening embryos for complex genetic traits called premature

Study of virtual embryos and real offspring suggests polygenic risk scores for IQ and height are poorly predictive

By **Jocelyn Kaiser**, in Houston, Texas

Two years ago, news headlines began to appear about a development that made many human geneticists uneasy. A U.S. company planned to offer a test for embryos created through in vitro fertilization (IVF) that screened the entire genome for DNA variants linked to cognitive ability, in order to help couples avoid having children with intellectual impairment. Many ethicists fear such multigene analyses could one day be used to screen embryos for desirable traits as well, such as tall stature or high IQ.

For those disturbed by the prospect, a study reported here last week at the annual meeting of the American Society of Human Genetics (ASHG) may come as a relief: For now, the strategy would not work very well.

Researchers, led by statistical geneticist Shai Carmi of the Hebrew University of Jerusalem, calculated exactly how much of a boost in IQ or height could be expected by scanning for relevant DNA markers in a batch of embryos and choosing those with the highest scores. The result: The gains would be slight, and prospective parents might even end up discarding their tallest or smartest potential offspring.

The work "is the first to empirically test the viability of screening embryos" for traits that are influenced by many genes, says sociologist and demographer Melinda Mills of the University of Oxford in the United Kingdom. Such embryo screening goes beyond today's testing for single-gene disorders and currently "isn't plausible," she concludes.

Such tests are based on a polygenic risk score, a tool for evaluating a person's likelihood of a disease or trait that has emerged over the past decade from genomic studies combing through variable DNA markers in many thousands of people. Although having any one DNA variant may barely raise the risk of, say, heart disease, adding up the effects from hundreds or thousands of these markers can generate a score that helps identify people at relatively high risk of common diseases. Some direct-to-consumer DNA testing companies have begun to give customers polygenic risk scores for diseases such as heart disease, breast cancer, and diabetes.

Testing embryos, however, is hugely controversial, because of both the scientific limitations of such polygenic scores and the prospect of designer babies. Undeterred, a company called Genomic Prediction last year began to offer to test cells plucked from an IVF embryo for millions of DNA markers to produce risk scores for some common diseases and for "intellectual disability" or low IQ. Co-founder Stephen Hsu, a physicist at Michigan State University in East Lansing who has branched into genomics, says that for now, the company is not returning genetic

Predicting genetically complex traits from the DNA of cells taken from an IVF embryo (above) isn't easy.

scores predicting high IQ because "society is not ready for it."

Still, testing embryos for desirable traits could be coming soon. Most people agree it's not a good idea, but there are no data, Carmi said at the ASHG meeting. To find out whether the strategy could work, his team created virtual genomes for potential embryos by combining the DNA profiles of "parents." One parental group included actual and randomly chosen pairs of men and women from 102 Ashkenazi Jewish couples with recorded heights and the other 919 randomly paired Greek men who had cognitive test scores. The team then calculated polygenic scores for the synthetic genomes to predict height or cognitive ability. For the five embryos typically generated

PHOTO: NATIONAL GEOGRAPHIC IMAGE COLLECTION/ALAMY STOCK PHOTO

in an IVF cycle, the theoretical height gain from selecting an Ashkenazi couple's embryo with the highest "tallness" score was about 2.5 centimeters (with a range of 1 to 6 centimeters). Selecting for the Greek virtual embryo with the highest IQ-favorable score brought a similarly limited gain in cognitive ability, just 2.5 IQ points (in a range from one to seven points), Carmi said.

His team also looked at the actual genomes of adult offspring in 28 large families (about 10 children on average). They found that for height, unknown environmental influences, which could include factors such as diet, and genes not represented in the polygenic score apparently overpowered the assessed genetic markers: In only seven of the 28 families was the sibling with the top score for height the tallest; in five families, the best scoring child was shorter than all siblings' average height. Although Carmi's team didn't have similar real-life data for IQ, University of Edinburgh population geneticist Peter Joshi expects that any intelligence polygenic score would be even more unreliable. "You might be wrong almost as often as you're right," Joshi says. (Carmi declined to discuss his study, which is online as a preprint and in press at a journal.)

An embryo screening test that uses a polygenic score to predict low IQ is likely to face the same limitations, says Joshi, who views such testing as unethical. Hsu, however, emphasizes that Genomic Prediction doesn't use its risk scores to screen for subtle IQ distinctions, but rather to avoid embryos with rare "outlier" DNA profiles for which scores predict a high risk of an IQ below 75, indicating intellectual disability.

Carmi's embryo study comes on top of other problems with polygenic tests. Most such scores have been derived from scanning DNA of people of European ancestry, making them of limited use for other ancestry groups. And recent studies have found they often aren't as predictive for older people, men, or people living in a certain location. Scores predicting behavioral traits such as a person's level of education are even more problematic, because these traits are strongly shaped by the family environment in which a child is raised.

But such polygenic scores, for IQ as well as diseases, are sure to improve as researchers look at genetic markers in larger, more diverse groups of people. "What about when they do become predictive?" an audience member asked during an ASHG discussion, wondering whether that would justify use of the scores. "That equation can change," responded Oxford statistical geneticist Alexander Young. The ethical debate over this brave new expansion of embryo screening is just beginning. ■

BIOMEDICINE

Prime editing promises to be a cut above CRISPR

New genome editor offers precision and versatility

By **Jon Cohen**

CRISPR, an extraordinarily powerful genome-editing tool invented in 2012, can still be clumsy. It sometimes changes genes it shouldn't, and it edits by hacking through both strands of DNA's double helix, leaving the cell to clean up the mess—shortcomings that limit its use in basic research and agriculture and pose safety risks in medicine. But a new entrant in the race to refine CRISPR promises to steer around some of its biggest faults. "It's a huge step in the right direction," chemist George Church, a CRISPR pioneer at Harvard University, says about the work, which appeared online this week in *Nature*.

This newfangled CRISPR, dubbed "prime editing," could make it possible to insert or delete specific sequences at genome targets with less collateral damage. "Prime editors offer more targeting flexibility and greater editing precision," says David Liu, a chemist at the Broad Institute in Cambridge, Massachusetts, whose lab led the new study and earlier invented a popular CRISPR refinement called base editing.

> **Prime editing "may become the way that disease-causing mutations are repaired."**
> **Fyodor Urnov**, Innovative Genomics Institute

Liu, his postdoc Andrew Anzalone, and co-workers tested variations of their prime editors on several human and mouse cells, performing more than 175 different edits. As a proof of principle, they created and then corrected the mutations that cause sickle cell anemia and Tay-Sachs disease, DNA aberrations that previous genome-editing systems either could not fix or only did so inefficiently. The edits occurred in a high percentage of cells and caused relatively few off-target changes. In its paper, the team claims the technology "in principle can correct about 89% of known pathogenic human genetic variants."

Most CRISPR systems rely on a molecular complex coupling a guide RNA that homes in on a specific location in the genome with an enzyme, Cas9, that cuts both strands of DNA. During the cell's efforts to reconnect the DNA, its repair machinery can introduce or delete nucleotides. Researchers can take advantage of the botched repair to knock out genes that, say, cause a disease. They can also hijack the inefficient repair process to add DNA—even an entire gene.

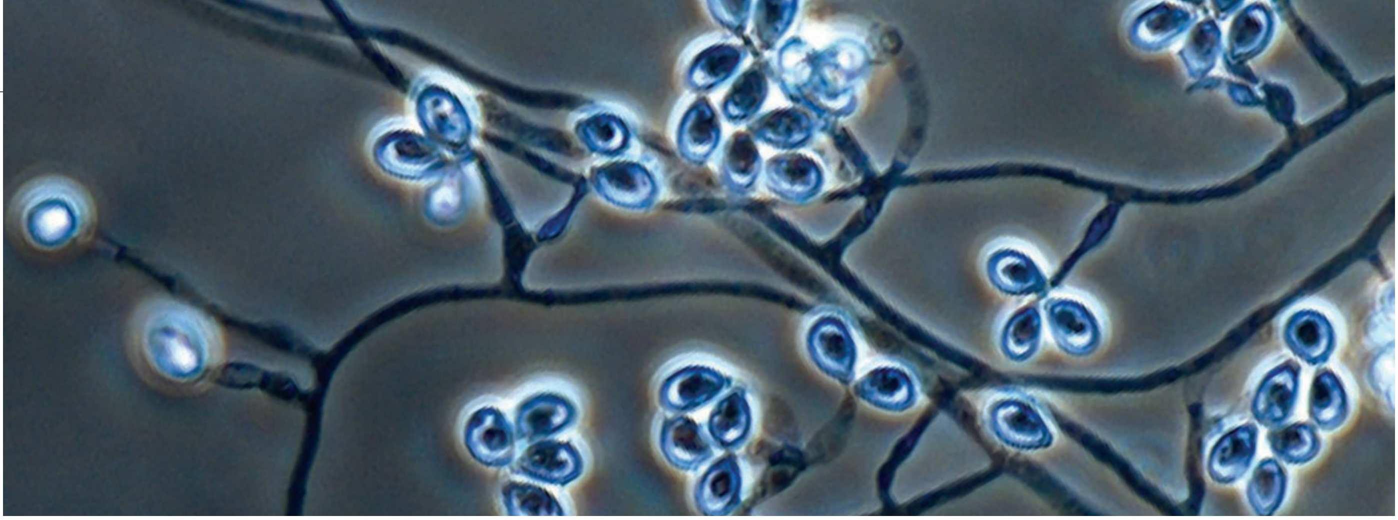
But double-stranded breaks are "genome vandalism," Church says. As the cell attempts to repair the break, it introduces insertions and deletions willy-nilly, sometimes creating unwanted—and even dangerous—mutations.

Liu's earlier handwork, base editing, does not cut the double-stranded DNA but instead uses the CRISPR targeting apparatus to shuttle an additional enzyme to a desired sequence, where it changes the identity of a single nucleotide. Many genetic traits and diseases are caused by a single nucleotide change, so base editing offers a powerful alternative for biotechnology and medicine. But the method has limitations, and it, too, often introduces off-target mutations.

Prime editing steers around shortcomings of both techniques by heavily modifying the Cas9 protein and the guide RNA. The altered Cas9 only "nicks" a single strand of the double helix, instead of cutting both. The new guide, called a pegRNA, contains an RNA template for a new DNA sequence, to be added to the genome at the target location. That requires a second protein, attached to Cas9: a reverse transcriptase enzyme, which can make a new DNA strand from the RNA template and insert it at the nicked site.

Liu, who has already formed a company around the new technology, Prime Medicine, stresses that to gain a place in the editing toolkit, it will have to prove robust and useful in many labs. Delivering the large construct of RNA and enzymes into living cells will also be difficult, and no one has yet shown it can work in an animal model.

Fyodor Urnov, scientific director at the Innovative Genomics Institute in Berkeley, California, reviewed the paper for *Nature* and says it brought "one of those 'yay, science!' kind of moments." Prime editing "well may become the way that disease-causing mutations are repaired," he says. But, he adds, it's too soon to be sure. The technique "just showed up this year." ■



Olorofim, now in phase II trials, has shown promise against *Lomentospora prolificans* (above), a fungus that mostly infects patients with weakened immune systems.

PHOTO: DAVID ELLIS

DRUG DEVELOPMENT

New drugs target growing threat of fatal fungi

Well-stocked pipeline could yield new tools to treat intractable infections

By **Kai Kupferschmidt**, *in Nice, France*

Martin Hoenigl, a specialist in fungal infections at the University of California, San Diego, sees half a dozen patients a year infected with a rare mold called *Lomentospora prolificans* that is resistant to all available antifungals. Doctors use combinations of two or three drugs in high doses to try to stop the infection, but usually to no avail. "Most of these patients die," Hoenigl says.

That may be about to change. Hoenigl just started to enroll patients in a phase II trial of a new drug called olorofim that holds promise against *Lomentospora*. It is one of several compounds now in trials that could give doctors much-needed new tools against these intractable infections. "We have got several drugs with new mechanisms of action and they look really good," says Tom Chiller, who heads the mycotic diseases branch at the U.S. Centers for Disease Control and Prevention in Atlanta. "It's exciting times," Hoenigl says.

Of the millions of fungal species on Earth, only a few dozen regularly cause human disease; fungi don't grow well at mammals' high body temperature, and a healthy human immune system is adept at dealing with the ones that do. But the HIV/AIDS epidemic and modern medical interventions such as chemotherapy and transplantation have led to a growing number of people with compromised or suppressed immune systems whose bodies can be overrun by a fungal invader. Symptoms vary widely depending on which organs are affected; a lung infection, for example, can lead to shortness of breath and cough. An estimated 1.5 million people die worldwide every year of invasive mycoses.

New drugs have been slow in coming because research funding has been scarce and investors prefer drugs against chronic diseases that patients take for life over ones that cure an infection. "There's not a huge incentive, there's not a big market," Chiller says. The urgency is growing, however, as cases have increased and once-treatable fungi are becoming resistant. About one-quarter of recent Indian isolates of *Candida auris*, a fungus that's on the rise globally, were resistant to two or more classes of antifungals, for instance. (For some other species, drugs have yet to be found.) And there are other problems: Only one class of drugs, the azoles, can be taken orally. The others must be injected, and many have side effects or interact with other drugs, a problem for patients who have a fungal infection on top of another illness.

At a medical mycology meeting here earlier this month, scientists presented promising phase II clinical data for rezafungin, a new member of an existing class called the echinocandins developed by Cidara Therapeutics in San Diego. Like other echinocandins, it acts by inhibiting synthesis of the polysaccharides that make up the fungal cell wall, but it has a much longer half-life, allowing it to be given once a week instead of daily. It might one day become the standard prophylaxis for patients receiving transplants, Hoenigl says.

Doctors are eager to have new classes of drugs—which have a new target or a new way to attack an existing target—because it allows them to try new combinations that may work synergistically and help avoid resistance. Olorofim, the compound Hoenigl is testing, is one example: It blocks the synthesis of pyrimidine, the precursor of DNA building blocks. Developed by U.K.-Austrian biotech F2G, olorofim can be taken orally and kills not only *Lomentospora*, but also *Scedosporium*, another rare and usually fatal mold infection. Further along is ibrexafungerp, developed by Scynexis in Jersey City, New Jersey, and now in a phase III trial. Like the echinocandins, it attacks the fungal cell wall but does so by latching onto another part of a key enzyme.

On these drugs' heels is fosmanogepix, a new broad-spectrum antifungal developed by Amplyx in San Diego that is now in phase II trials. Last month, the U.S. Food and Drug Administration gave it "fast-track" status, an expedited review procedure for urgently needed drugs.

Because fungal diseases are relatively rare and diagnosis is difficult, the current clinical trials are small. But Oliver Cornely, an expert on fungal infections at the University of Cologne in Germany who is involved in trials of all four drug candidates, hopes that even a limited amount of positive data will persuade regulators to approve these drugs, which address fatal diseases with few treatment options. "If this were HIV, we would see people protesting in front of the convention center for these drugs to become available."

Chiller says new diagnostics are also needed. Fungi often lead to unspecific symptoms and are hard to culture, and rapid, accurate diagnostic tests are scarce. That means patients may die without ever receiving a proper diagnosis. For the new drugs to have their biggest impact, that has to change, Chiller says. "If you can't diagnose it, it doesn't exist." ■

Students shaped stone tools at an all-female field camp run by the University of Cape Town in South Africa, where they also workshopped a code of conduct for fieldwork.

PHOTO: ROBYN WALKER

SCIENTIFIC COMMUNITY

Codes of conduct aim to curb harassment at field sites

Guides spell out behavior and navigate cultural differences

By **Linda Nordling**

As an African female archaeology student, Mary* is accustomed to feeling like an outsider. She has attended five field schools in Africa where most students and staff were white foreigners. Some of those foreigners expressed surprise that she speaks English and has a degree in archaeology. And some commented on her body in a way that she says "fetishizes black women."

One field school had a sexual harassment policy participants had to sign, but it referred to Title IX, the 1972 U.S. law banning institutions that receive federal funds from discriminating on the basis of sex. For Mary, who hails from an African university, Title IX and any support offered through its offices in a U.S. institution felt meaningless. After reading the code, she felt no protection at all. "As a black woman in paleoscience, I feel more needs to be done."

Stronger codes of conduct could help make international sites safe for everyone, says biological anthropologist Rebecca Ackermann, deputy dean at the science faculty of the University of Cape Town (UCT) in South Africa. She's spearheading a model code, part of a document guiding conduct at science departments at UCT, to be announced this month. UCT's effort adds to a growing movement to adopt rules of conduct at field sites worldwide, aimed at prejudices and cultural differences.

Studies of fieldwork experiences and high-profile cases of sexual misconduct during fieldwork, including at African-based field schools (*Science*, 13 October 2017, p. 162 and 12 February 2016, p. 652), have boosted awareness about the dangers of remote field sites, where scientists from many countries live and work together. But most fieldwork policies around the world are "very much entrenched in an American definition and reporting structure," Ackermann says.

UCT's new fieldwork code was designed with South African needs in mind, Ackermann says. The code bans inappropriate jokes, persistent one-sided flirting, or objectifying comments on somebody's dress, looks, or sexuality. Clauses ban alcohol and relationships between scientists and locals. The code also prohibits types of discrimination relevant to South Africans, including those based on race or HIV status.

The code is a model, not mandatory, and individual site leaders are free to adjust provisions like the alcohol ban. But the code spells out that violations will result in "immediate removal" from the project.

The code "sets the tone for more professional and inclusive conduct," says Jemma Finch, a paleoecologist at the University of KwaZulu-Natal in Pietermaritzburg, South Africa, who applauds the effort.

Although few universities have drawn up templates like this, a growing number of field projects are writing their own, navigating diverse cultural norms to do so. For example, Tuanan Orangutan Research Station in Indonesia, a majority Muslim country, has had a code of conduct since 2016, says primatologist Erin Vogel from Rutgers University in New Brunswick, New Jersey, who co-directs the station with Suci Utami-Atmoko from Jakarta's National University. Some years ago, an intimate relationship between unmarried people at the camp became common knowledge and caused conflict with the local community. After a shaman led a 24-hour cleansing ceremony, Tuanan developed the written code, which applies to everyone in camp.

The code prohibits "following people around camp when asked to stop or telling them you want to be with them in a sexual way." It also bans behaviors deemed inappropriate in the region, including intimate relationships between unmarried people. Violators face immediate termination.

There's been some pushback from people whose cultures don't prohibit consensual unmarried relationships, Vogel says. Her message: "Go to Bali and do what you want. … Just don't do it in camp." The code seems to work, she says. "We haven't had any issues in camp for years."

Mary and others caution that codes might not be enough to change behavior. She'd like to see sensitivity training for staff and students, too, in part to challenge patriarchal African cultures. "Some typically African males still hold the view that sexual harassment is a Western concept."

Ackermann acknowledges the problem. If local cultural norms clash with women's rights, she says, a field site can send the message that "in your culture it might be the norm, but in this culture of science it's not acceptable."

The situation becomes even more challenging when local laws as well as customs are discriminatory, for example regarding LGBTQ identity. Thirty-four African countries outlaw homosexual activity. For Silindo*, a South African geologist who is gay, planning fieldwork in those countries has been "very overwhelming."

For him, the international atmosphere in many field sites helps him feel protected. "It is good to have these conversations before you go to a place like this," he says. "We need to create these [respectful] spaces—they don't just happen overnight." ■

*Names have been modified to protect identities. Linda Nordling is a journalist in Cape Town, South Africa.

PALEONTOLOGY

How life blossomed after the dinosaurs died

Plants and mammals diversified together, blow-by-blow record shows

By **Elizabeth Pennisi**

In 2014, when Ian Miller and Tyler Lyson first visited Corral Bluffs, a fossil site 100 kilometers south of the Denver Museum of Nature & Science where they work, Lyson was not impressed by the few vertebrate fossils he saw. But on a return trip later that year, he split open small boulders called concretions—and found dozens of skulls. Now, he, Miller, and their colleagues have combined the site's trove of plant and animal fossils with a detailed chronology of the rock layers to tell a momentous story: how life recovered from the asteroid impact that killed off the dinosaurs 66 million years ago.

Plants and animals came back much faster than thought, with plants spurring mammals to diversify, the team reports online in *Science* this week. "They get almost the whole picture, which is quite exciting," says functional anatomist Amy Chew of Brown University. "This high-resolution integrated record really tells us what's going on."

When the asteroid slammed into Earth, it wiped out 75% of living species, including any mammal much larger than a rat. Half the plant species died out. With the great dinosaurs gone, mammals expanded, and the new study traces that process in exquisite detail.

Most fossil sites from after the impact have gaps, but sediment accumulated nearly continuously for 1 million years on the flood plain that is now the Corral Bluffs site. So the site preserves a full record of ancient life and the environment.

Such sites can be hard to date. But Miller, a paleobotanist, and his colleagues collected 37,000 grains of pollen and spores, which revealed a clear marker of the asteroid impact: a surge in the growth of ferns, which thrive in disturbed environments. The site also includes two layers of ash from nearby volcanoes. Volcanic ash includes radioactive minerals whose decay can be used as a precise geochronological clock, providing two time markers. The known flips in Earth's magnetic poles, which some minerals in the layers had recorded, add detail to the chronology. "They have a very strong geochronological framework," says David Fastovsky, a paleontologist at the University of Rhode Island in Kingston.

The record confirms the devastation wrought by the impact. Racoon-size mammal species had swarmed the site before the catastrophe, but for 1000 years afterward just a few furry creatures no bigger than 600-gram rats roamed a ferny world where flowering plants, with their nutritious seeds and fruits, were scarce.

By 100,000 years later, twice as many mammal species roamed, and they were back to raccoon size. These critters foraged in the palm forests that replaced the ferns. "It's a world that's coming back from complete and utter devastation," Miller says.

Over the next 200,000 years, what he calls the "palm period" gave way to the "pecan pie" period, when walnutlike plants arose. New mammals evolved to take advantage of the nutritious seeds. Mammal diversity increased threefold, and the biggest of the new species reached 25 kilograms—beaver size.

After about 700,000 years, legumes showed up; their fossil pea pods are North America's oldest discovered to date. Pea and bean species from the "protein bar period" provided protein-rich meals that further boosted mammalian size and diversity, Lyson says. Mammals topped 50 kilograms—a 100-fold increase over those that survived the asteroid. The forests, too, had recovered. "The biggest message is how fast the recovery was … and how closely the vegetation and fauna are tied together," says Vivi Vajda, a paleobiologist at the Swedish Museum of Natural History in Stockholm.

The team also classified 6000 leaves, counting how many species at each time interval had smooth or toothed edges. Smooth-edged species are more common in hot climates. The team concluded that the site underwent three warming periods. They estimate that the first, just after the impact, saw temperatures rise about 5°C, agreeing with earlier work. This period coincides with the massive volcanic eruptions of India's Deccan Traps, which could have warmed Earth by belching carbon dioxide (*Science*, 22 February, p. 862 and p. 866).

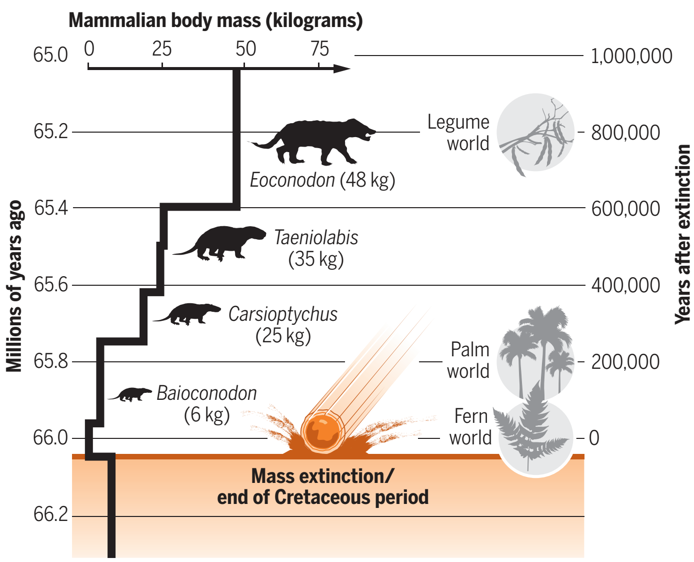
"At each warming period you see a change in the plant community and subsequently, changes in the mammals," says Lyson, who thinks temperature drove the stepwise recovery.

Vajda thinks no matter what happened to temperature and plant life, the loss of dinosaurs alone might have opened the door to bigger, more diverse mammals. But Jukka Jernvall, an evolutionary biologist at the University of Helsinki, says the team's analysis of ancient ecosystems shows just how the recovery unfolded. "We are starting to get the time and spatial resolution to reconstruct the environment and what happened in a way that can be linked to ecological processes."

The record also holds a sobering message about the future, and how quickly ecosystems might recover from ongoing, human-driven extinctions. Even a recovery that geologists call "fast" took hundreds of thousands of years, and the world was never the same. "A very dramatic resetting of the ecosystem could be in our future," Chew says. ■

A stepwise recovery

After an asteroid wiped out much of life on Earth, mammals—responding to changes in plants—grew in size and diversity surprisingly quickly.



GRAPHIC: DENVER MUSEUM OF NATURE & SCIENCE. ADAPTED BY C. BICKEL/SCIENCE

SPACE SCIENCE

Europe dreams big for future space missions

Scientists debate themes vying for billion-euro mission slots in Voyage 2050 program

By **Daniel Clery**

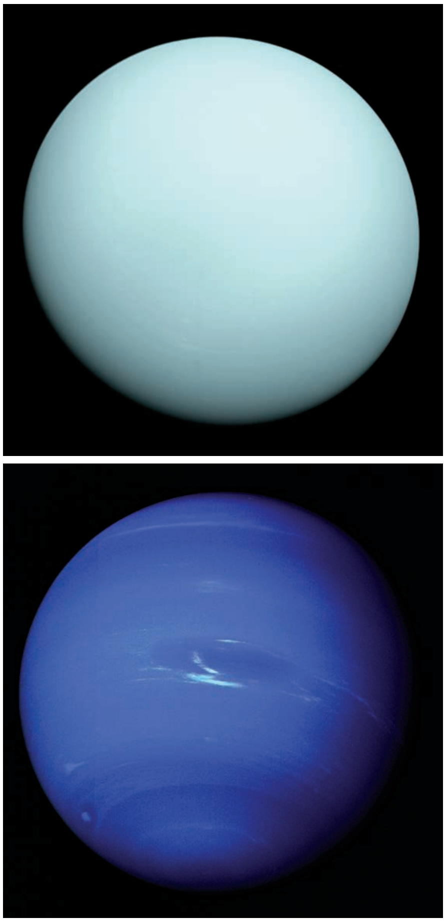
European space scientists' most ambitious dreams are about to meet reality. Earlier this year, the European Space Agency (ESA) asked a committee to invite researchers to pitch ideas for Voyage 2050, a 20-year program of missions beginning in 2035. "We want to know what their dreams are," says astronomer Linda Tacconi of the Max Planck Institute for Extraterrestrial Physics in Garching, Germany, who chairs the committee. It received almost 100 proposals, including scientific quests that would require an orbiting radio telescope the size of several city blocks, a fleet of formation-flying telescopes to study exoplanets, and a robotic odyssey to bring back samples from Venus. Next week, scientists will gather in Madrid for a workshop to debate the merits of the different science themes, after which the committee will spend about 1 year choosing which ones to recommend to ESA for the program.

Günther Hasinger, ESA's director of science, says the agency is looking to lead three large billion-euro missions, and five or six medium missions costing less than half a billion euros—an agenda similar to its current science program, called Cosmic Vision. Voyage 2050 will also include a series of "fast" missions, with budgets less than €150 million, launching every 3 years. This year, ESA picked the first of those, the Comet Interceptor, which after launch in 2028 would park itself beyond the Moon and wait for a pristine comet to enter the Solar System.

Tacconi's committee will settle on the themes for the three large missions and a menu of about a dozen for medium missions. Calls for specific mission proposals to fit the themes will come later. Hasinger says the long lead time will help ESA get a head start on technology development for the missions.

Many of the proposals submitted for the workshop call for exoplanet science, Hasinger says. One envisions a few mid-infrared telescopes orbiting in formation about 150 meters apart and combining their light to get sharp views of Earth-like exoplanets. Another calls for a parasol-shaped starshade orbiting in space, which would create brief eclipses of distant stars so that ground-based telescopes can discern the faint light of planets around them.

Leon Koopmans of the University of Groningen's Kapteyn Astronomical Institute in the Netherlands heads a team that wants to probe the universe's dark ages, before the light of the first stars. Because Earth's atmosphere blocks radio signals from that era, the team envisions a giant orbiting radio antenna with a collecting area of 1 square kilometer and, later, an array of antennas on the radio-quiet far side of the Moon. "It's been treated as an exciting and original idea. But there is stiff competition," Koopmans says.



The European Space Agency wants to partner with NASA on a mission to Uranus (top) and Neptune, last visited by Voyager 2 in the 1980s.

PHOTOS: (TOP TO BOTTOM) NASA/JPL; NASA

Other ambitious proposals have targets closer to home. One would retrieve samples from the surface of Venus through a complex choreography of two orbiters, multiple landers, a balloon high in the atmosphere, and a return rocket.

Constraining those ambitions is ESA's science budget, which all 22 member countries must contribute to in line with their gross domestic product. It has been flat for the past couple of decades. As inflation erodes its value, the gaps between missions have grown longer. Hasinger points out that no medium-size science mission has launched in more than 10 years.

Next month, government ministers from the 22 member nations will meet in Seville, Spain, to hash out the agency's budget for the next 5 years or so. The science directorate is asking for a funding boost of between 10% and 20%. "This is absolutely necessary. A flat budget is not healthy," says Athena Coustenis, a planetary scientist at the Paris Observatory and chair of the European Space Sciences Committee, an advisory body. "There's almost unanimous support that there has to be an increase, but the devil is in the detail," Hasinger says. "We may not get as much as we asked for."

Hasinger has other items on his wish list for Seville. He will also ask ministers for extra funding for Cosmic Vision, so ESA can launch one of its large missions, the Laser Interferometer Space Antenna, a gravitational wave detector, earlier than its scheduled 2034 date. That way, it can fly at the same time as another large ESA mission, dubbed Athena, an x-ray observatory due for launch in 2031. "They both study the same phenomena but with different eyes and ears," Hasinger says, explaining that the cosmic collisions that generate gravitational waves can also unleash blasts of x-rays.

ESA also wants to join NASA in sending a probe to Uranus and Neptune, the first visit to the ice giants since Voyager 2 in the 1980s. Around 2030, the planets will be aligned with Jupiter so it can act as a gravitational slingshot to fling a spacecraft onward. This maneuver saves so much fuel that the spacecraft can be five times bigger than without the boost, Hasinger says. Final approval for the ice giants mission will likely have to wait until the next ministerial meeting after Seville, because the NASA team is also seeking support, from the U.S. space science community. ■

FEATURES

GROWING PAINS

Ecologist Thomas Crowther is having a bumpy rise to prominence

By **Gabriel Popkin**

It's been a remarkable year for ecologist Thomas Crowther—for better and worse. The 33-year-old researcher, an assistant professor at ETH Zurich in Switzerland, has co-authored six papers in *Nature* and *Science*. He's addressed members of Germany's parliament and is advising major companies and nonprofits on environmental issues. Journalists have produced hundreds of stories about his laboratory's innovative work with huge ecological data sets, including a headline-grabbing estimate, published in *Science*, that planting trees on hundreds of millions of hectares could dramatically slow climate change. Last month, *Nature* published a lengthy profile of Crowther, noting his rise "from struggling student to steward of a 30-strong team," supported by a hefty foundation grant "that should keep his lab going for 13 years."

But Crowther's rapid rise to prominence has also catalyzed a strong and sometimes personal backlash. Colleagues have called him out for what they view as preening for the press (his lab has its own public relations staff) and for publishing flashy but flawed science. Critics panned the tree-planting paper, with one calling it "shockingly bad," and last week, *Science* published six lengthy critiques signed by more than 50 scientists. And the *Nature* profile, in which Crowther appeared to belittle traditional ecological research, incited a Twitter pile-on. Some users accused Crowther of being a poster boy for white male privilege in science, with one deriding his lab as "a huge waste of money."

The attacks have left Crowther and members of his laboratory feeling bewildered, even depressed. But the attention has also reinforced their feeling that they are doing meaningful, precedent-setting, and world-changing work. "Even though this has been absolutely the worst 3 months of my career," Crowther says, "it's also been the 3 months when I've realized the dream of actually having our science be useful."

Few doubt that Crowther's efforts to generate global ecological numbers are having an

PHOTO: FREDERIK VAN DEN BERG

Thomas Crowther's work, especially on trees, has produced high-profile publications, but also criticism.

impact. He's "very savvy in picking problems that are going to have answers that people are going to want to know," says ecologist Jonathan Levine of Princeton University. "There's sort of a buzz around Tom."

CROWTHER APPEARS to be all limbs, tall and lanky. He sometimes dons a three-piece suit for talks but favors less formal garb in his lab: for example, a ball cap, faded T-shirt, flip-flops, and hoodie. He gesticulates energetically and often speaks in hyperbole, with a healthy dose of expletives. Everyone in his lab is the most brilliant scientist or most incredible programmer he's ever met. During one recent talk he presented "the graph that gave me the most joy in my entire career."

Crowther's accent hints at his multinational upbringing. He was born in Namibia and spent his early years in South Africa, before his family moved to the coastal town of Prestatyn, U.K. His parents are doctors and his childhood was comfortable, but dyslexia made academics hard. He became known as a troublemaker and poured his energy into sports, rising to the apex of youth tennis in the United Kingdom before, he says, a lack of focus cost him his funding. At the same time, he grew fascinated with nature, devouring documentaries narrated by naturalist David Attenborough and developing a passion for catching lizards and snakes.

Crowther enrolled at Cardiff University, but did poorly until a soil ecology professor, Hefin Jones, saw his potential and, over a beer, encouraged Crowther to try harder at science. As he had done with sports, Crowther threw himself into experiments: growing wood-decomposing fungi in petri dishes, then examining the consequences of adding fungus-eating insects to the mix.

Once, Jones recalls, Crowther insisted on submitting a paper to *Ecology Letters*, a prominent journal, even though Jones thought the result was trivial. The journal accepted it with few revisions. The episode proved Crowther's "willingness to actually disagree with his seniors and … argue his case," Jones says. By the time Crowther completed his Ph.D. at Cardiff in 2012, he had published in several top ecology journals.

In 2012, Crowther moved to Yale University to work with soil specialist Mark Bradford. Soon after he arrived, a German high school student, Felix Finkbeiner, emailed Crowther's roommate with a question: How many trees are on Earth? Finkbeiner was curious because he ran a nonprofit that was working with the United Nations to encourage people to plant 1 billion trees.

"Why would a scientist care?" ecologist Daniel Maynard, then a grad student at Yale, recalls thinking. But Crowther embraced the challenge of amassing the data. He began to ask tree researchers to share their numbers. Some did; others refused. Then, Crowther discovered that many governments had massive data sets of forest measurements. Soon he had 400,000 data points, enough to correlate tree density with variables such as rainfall and estimate tree numbers in places with sparse data, including much of Africa and Asia. In September 2015, Crowther and colleagues—including the once-dubious Maynard, who is now in Crowther's lab—reported in *Nature* that the planet is home to about 3 trillion trees.

Journalists began to call. Attenborough quoted the paper. ("The best moment of my career," Crowther says.) The Billion Tree Campaign became the Trillion Tree Campaign.

Scientists had mixed reactions. Some felt Crowther had neglected more meaningful metrics, such as species distribution and function. Robin Chazdon, a forest ecologist at the University of Connecticut in Storrs, recalls thinking: "The number of trees is just kind of a flashy way of getting attention."

By the time the *Nature* paper appeared, Crowther had moved to the Netherlands Institute of Ecology (NIOO) in Wageningen. He says he learned two things from the experience. One was that he could pull off—and publish in a prestigious journal—a credible global analysis of a single ecological variable. "People said: 'How can you be confident enough to make claims about the whole world?'" he recalls. "I said: 'It's identical to my petri dish. … Being bold enough to make statements about petri dishes is the same as being bold enough to make statements about the globe."

The other lesson was that relatively simple studies, when clearly communicated, could excite the media and the public. "There's loads of genius scientists way smarter than me who are not willing to communicate their science, so it doesn't get the attention it deserves," Crowther says. "Science isn't about finding absolute perfection; it's about moving the conversation forward."

CROWTHER APPLIED those lessons to his next big project. This time, his spotlight was on soil, which is often overlooked yet is vital for growing food—and storing carbon. In particular, many researchers want to know how global warming might change the activity of soil microbes, which play a major role in determining how much carbon is stored or released by soil. Researchers have launched numerous experiments in which they artificially warm soil in different places and measure what happens.

As he did for tree numbers, Crowther drew on other researchers' data. He plotted the results of 49 experiments and found a trend: Warming soils lost carbon, and the more carbon a site held, the more it lost. "It's the simplest thing ever. It's just that I dared to apply the concept at a global scale," he says. "I would never have been bold enough to try, until the tree paper."

Some of Crowther's co-authors wanted to emphasize the uncertainty in the results. But Crowther prevailed with a simpler message: A 1°C rise in global temperature could cause microbes to release about 55 billion tons of soil carbon by 2050, accelerating global warming by up to 17%. The result would be, he said, like adding emissions from an additional United States to the atmosphere.

The work, published in *Nature* just 15 months after the 3 trillion trees paper, received extensive coverage. Crowther also received an invitation to apply for funding from DOB Ecology, a family foundation in Veessen, Netherlands. The foundation rejected Crowther's first proposal as not ambitious enough. But it liked his second pitch, which called for creating a team including programmers, remote sensing experts, mod-

Published by AAAS

elers, and communications specialists to conduct and publicize 13 global-scale studies with practical applications. In October 2017, 1 month after Crowther joined ETH Zurich, he received a $2.7 million grant from DOB Ecology, with a promise of an additional $15 million if he meets certain targets.

Overnight, Crowther had more money than most in his field. "I've never heard of an assistant professor having this degree of financial support, at least in ecology," says Levine, who was at ETH at the time. Crowther says he was embarrassed by an ETH press release crowing about the "huge" grant. "It's horrible, it's divisive. … I did freak out for a little bit," he says. "The reason I've had academic success is that I've never had stress. … Then I got into this, and I'm like: 'Oh my God, I'm responsible for 20 people.'"

NOW, 2 YEARS LATER, Crowther's team of about 30 occupies a hallway of ETH's environmental sciences building. The space has the feel of a startup, with lounge furniture, espresso cups, and whiteboards strewn about. You will often find Crowther at a standing desk in an office shared with five others. He declined to occupy an office next door that was offered to him. "I get so sad and bored just sitting in a room by myself," he says. Instead, it's become a game room, equipped with a table that hosts frequent bouts of "smash," a chaotic form of table tennis that Crowther invented in which players bounce the ball off walls, the floor, their bodies, and even visiting journalists.

Crowther's penchant for games and camaraderie extends beyond the office. One Friday night, he led a lab pub crawl that lasted into the early morning. At a weekend picnic, a restless Crowther searched for the optimal angle for swinging on a rope from a tree while his labmates drank beer. "Zurich was a classy place until the Crowther lab showed up," a colleague joked.

Crowther has landed like a cannonball in Switzerland's traditionally calm academic pond. ETH officials balked when Crowther sought to hire his own publicity team. He did it anyway. And ETH faculty have chastised him for some careless remarks to the press, says environmental scientist Nicolas Gruber, an ETH colleague. But he says Crowther's "ambition level and translating that ambition to action [has been] quite amazing."

This year alone, Crowther's lab has helped produce the first-ever global distribution maps of soil-dwelling fungal groups and of tiny worms called nematodes, both in *Nature*, and a global map of earthworm distribution, in this issue of *Science* (p. 480). The results will help bolster climate and ecological models, Crowther says. His team has published other work on how climates in major cities could shift by 2050.

Some of the lab's work is more hands on. Graduate student Julia Maschler is examining how climate change might affect tree growth and wood decomposition. In another study, Crowther's playfulness comes through. He speculated that exposure to sound would make fungi grow faster. Maynard, now a lead scientist in his lab, bet it wouldn't. To settle the issue, they grew fungi in chambers rigged for sound. Crowther won.

IT IS THE CROWTHER LAB'S tree maps, however, that keep stealing the show—and engulfing him in controversy. Ecologist Jean-François Bastin, a postdoc in the lab, led an effort to determine how much unforested land on Earth could support trees. It was prompted, in part, by a practical problem: Tropical nations have committed to establishing forests across an area larger than India by 2030, but many countries don't know how much land is actually suitable.

The researchers, including computing specialist Devin Routh, built models that drew on variables such as soil type to estimate how much land could be forested, and how much carbon dioxide those trees might suck from the atmosphere. In July, they reported in *Science* (5 July, p. 76) that new trees could be planted across 900 million hectares of land, where they would absorb two-thirds of the carbon dioxide humans have added to the atmosphere in the industrial era. Tree planting, the authors wrote, is "our most effective climate change solution."

> **"We cannot be that bold. We get bogged down in bureaucracy."**
> **Julian Fox**, Food and Agriculture Organization of the United Nations

The media embraced that message. English-language outlets ran some 700 stories on the paper, according to a firm Crowther hired to promote it. Climate advocate Greta Thunberg and former Vice President Al Gore, among many others, cited the result.

But many scientists objected. One of the bluntest critics was Simon Lewis, a forest ecologist at the University of Leeds in the United Kingdom, who had previously collaborated with Crowther. "Science Magazine have lost their critical faculties," he tweeted, calling the paper "shockingly bad." In particular, Lewis faulted the carbon removal calculation, arguing that Crowther's team had overstated, by a factor of two, the proportion of human-emitted carbon that trees could absorb. "That's not how the global carbon cycle works," he wrote.

Crowther and Bastin were taken aback. The attacks "stopped all work in the lab for a week," Crowther recalls.

Last week, the conflict got a full airing in *Science*, which published six submissions from critics (including Lewis), together with one supportive comment and responses from Crowther's team (*Science*, 18 October, p. 315). In addition to lambasting the July paper for miscalculating carbon storage, the critics argued it favored converting grasslands and wetlands to forests and ignored how trees might affect water supplies and temperatures. "The claim that global tree restoration is our most effective climate change solution is simply incorrect scientifically and dangerously misleading," one group wrote.

Crowther's team revised the paper abstract to recharacterize tree planting as just "one of the most effective carbon drawdown solutions" and emphasize that reducing carbon emissions is critical. But they vigorously challenged other objections. Disagreements about carbon storage calculations were not the result of errors, but different definitions of "forest" and confusion about their methods, they wrote. And they denied promoting conversion of grasslands and wetlands. "All we're showing is where trees can grow," Crowther says. "It's not got any agenda."

Ecologist Jonathan Foley, executive director of Project Drawdown in San Francisco, California, which ranks climate solutions, says his group's analysis confirms Crowther's numbers. And Julian Fox, national forest monitoring team leader at the Food and Agriculture Organization of the United Nations in Rome, credits Crowther's team for showing that some countries have made unrealistic forestation promises, whereas others could be more ambitious. His team had launched a similar study, but it realized Crowther's team could deliver harder hitting results. "We cannot be that bold," he says. "We get bogged down in bureaucracy."

Amid the dueling, even scientists who have been critical of Crowther's work say his messaging has helped advance meaningful conversations. Chazdon says the *Science* paper "overinflated" the potential for carbon removal, but that Crowther is "really raising the profile of forests" in discussions of climate solutions. Although she had been wary of working with Crowther in the past, Chazdon has agreed to share some of her data with him. "He's matured," she says.

THIS SUMMER, a stream of corporate social responsibility officers, wealthy donors, and nonprofit directors filed into a domed conference room on the ETH campus for a Crowther lab "launch" event. Standing be-

Members of the Crowther lab take the stage at an event held this summer for corporate social responsibility officers, donors, and nonprofit officials in Zurich, Switzerland.

fore a curved backdrop depicting a forest, Crowther breezed through highlights of his work on how nature can help fight climate change. The feel was more TED Talk than data dump, and he ended his remarks with an upbeat message that drew applause. "Until recently, fighting against climate change has always been about things we've got to stop—things we shouldn't do," he said. "Now, we've got positive actions we can take."

It's a message that fits today's corporate zeitgeist. Firms are awakening to climate change threats and are eager to support solutions such as tree planting. They've embraced Crowther, recruiting him for advisory boards and consulting gigs. (Crowther says he gets no money from corporations.) But outside of conferences where scientists and corporate officers swap sustainability pledges and business cards, tree-planting projects face a tougher reality. Many planted trees don't live long enough to fight climate change, says Finkbeiner, who is now at Crowther's lab. Even when trees thrive, they are often cut too quickly to affect carbon levels.

Another problem is that forest restorers typically lack scientific expertise in augmenting soil microbial communities to sustain and optimize tree growth, says Louise Vet, a biologist at NIOO who chairs DOB Ecology's advisory council. Now that Crowther has made a splash with global papers, she and the foundation directors are pushing him to focus on site-specific research that will help DOB Ecology's 11 restoration projects succeed. "That's what I'm expecting from Tom—that he bring in some of the knowledge that can be used in practice by these partners."

Crowther is confident he can do that. He is hiring for a team, dubbed RESTOR, that will evaluate forest restoration projects throughout the tropics and help funnel donor money to well-designed efforts. "We need to go all the way there to making it happen," Crowther says. "We can't just rely on people taking our science and doing something with it."

EVEN AS CROWTHER sought to put the tree paper controversy behind him, a new storm hit. Soon after the 20 September publication of the *Nature* profile, social media erupted. Some researchers were upset by quotes in which Crowther appeared to dismiss the value of observational and field studies. "I can say, 'That bird is flying weirdly'—that's not science; that's what most of ecology is at the moment. It's natural history," he is quoted as saying. Collecting field data is "the simplest thing ever," he suggested. Others reacted to the story's illustrations, including a video of Crowther playing smash with three white male colleagues, which some saw as highlighting a white scientific patriarchy.

One of the harshest reactions came from Katherine Gould, a biologist at Pasadena City College and Moorpark College in California. Crowther was a "boy wonder who is funded by a rich family," Gould tweeted, referring to his lab with a profanity and calling it a waste of money. Gould's thread received some 500 retweets and 1500 likes, and sent Crowther and his lab into damage-control mode.

"It was horrendous, I was depressed," Crowther recalls. He even considered leaving academia. But he was buoyed by supportive tweets from several female lab members. "I'm a happy female PhD student supervised by an ever engaged and enthusiastic" Crowther, tweeted Iris Hordijk, who studies tree demography at ETH. "Coming from a mechanical engineering lab, I've never seen so many women scientists in one lab," tweeted Tanja Koch, who recently completed a master's degree with Crowther. "And believe me the girls love smashing game as well," she added.

Crowther himself tweeted, "Just to clarify, I truly love natural history. That's why I got into ecology. I would never intend to criticize the incredible ecologists who do amazing work …" Some of his comments had been taken out of context, he protested, but apologized for any miscommunication. The flap faded, and Gould apologized for her initial characterization of Crowther and his lab. But she continued to take Crowther to task for a "colonial and paternalistic" approach to research and his lab's lack of racial diversity.

Crowther has acknowledged that unearned privilege that comes from being a tall, charismatic, middle-class white man has likely played a role in his rapid ascent. "I've been really fortunate," he said this summer, before the *Nature* profile appeared. "I do believe that my privilege has made me more persuasive … whether it's writing papers or encouraging a really good postdoc to come."

Crowther also says he's "seriously worried about" the fact that his senior lab members are all men, and says he is emphasizing efforts to recruit women. He recently hired a female postdoc to join the lab's leadership, part of fulfilling a vow to fill at least four of the lab's next five positions with women.

Whether such efforts will satisfy critics remains to be seen. But the back-to-back firestorms have convinced Crowther that he might need to rethink his exuberant but informal style, which has gotten him so far so fast. And he has learned twice over a lesson that every prominent person learns sooner or later: that fame and fortune can cut both ways. ■

Gabriel Popkin is a journalist based in Mount Rainier, Maryland.

INSIGHTS

PERSPECTIVES

PHOTO: MYUNG J. CHUN/LOS ANGELES TIMES VIA GETTY IMAGES

PLANT BIOLOGY

Plant hydraulics and agrichemical genomics

A next-generation synthetic hormone is developed to understand water use in plants

By **George N. Phillips Jr.**[1] and **Michael R. Sussman**[2]

A challenge of continued efforts to increase crop yields to feed the expanding population is that humans are competing with crop plants for clean water. Treatment of wastewater is required not only for drinking, but also to water crops. For example, when wastewater is used on crops, some vegetables accumulate synthetic contaminants, including antibiotics and psychoactive drugs arising in part from personal care products and pharmaceuticals disposed in drains (*1*). Ensuring that crop plants grow efficiently even when clean water is in short supply is crucial to secure food production. On page 446 of this issue, Vaidya *et al.* (*2*), using prior knowledge of the structure of a family of plant proteins that play a key role in drought tolerance, introduce a new generation of synthetic compounds that may lead to reduced water use in agriculture.

Although animals and plants both need clean water, plants use water as a hydraulic system much like bones and muscles—that is, for strength in maintaining an erect structure, and for movement of root and shoot tissues in response to gravity and sunlight. Thus, one of the most important water-associated properties in plants is turgor pressure, the force of the plasma

The power of turgor pressure:
Tree roots can burst through pavement.

membrane against the thick cellulose cell wall. In plants, turgor pressure is high when the cells have high concentrations of osmolytes, such as K^+ ions. When osmolyte concentrations are low, the cells become flaccid. Because of the semi-rigid cell wall, turgor pressure in plants is a formidable force necessary for plant growth and development. A graphic example of this is the observation of tree roots disrupting pavement (see the photo).

Although the hormone abscisic acid (ABA) affects water usage in all plant cells, there are two cell types in which it plays especially important roles. After fertilization of the ovule by pollen, ABA is responsible for regulating the water content of the developing embryos as they are desiccated naturally on their way to becoming dry seeds. This hormone plays an equally important role in regulating the water content of guard cells in leaves. Guard cells form pores (called stomata) in leaves, and these holes are the main pathway through which water is lost and carbon dioxide (used for photosynthesis) enters the plant. The state of inflation of guard cells, determined through turgor pressure, regulates the size of the pore that forms in the space between them. ABA and guard cells have thus been a major focus of basic and translational plant research for many decades.

Unlike animals, plants use proton gradients to regulate water pressure at the plasma membrane. A protein transporter that pumps protons (H^+) out of plant cells can generate much higher membrane potentials than are found in animals. This provides a huge driving force for the major osmolyte, K^+, to become concentrated inside the cytoplasm via potassium channels (*3*). In touch-sensitive plants, such as *Mimosa pudica* and the Venus flytrap (*Dionaea muscipula*), the rapid movement of plant parts generated by touch is driven by an action potential that changes at the millisecond time scale and leads to rapid collapse of the membrane potential and loss of intracellular K^+. This in turn leads to rapid loss of water and, consequently, loss of turgor. In guard cells, instead of rapid mechanosensitivity-induced turgor changes, slower hormone-mediated turgor pressure changes occur, which result in opening or closing of the stomatal pore. Therefore, engineering ABA-like compounds that specifically alter guard cell function is a key aspect of current agricultural research to lower water loss via stomatal transpiration from plants.

In a series of landmark discoveries in 2009 arising from an analysis of *Arabidopsis thaliana* mutants together with in vitro reconstitution and x-ray crystallography, the critical early events in ABA perception were elucidated (*4–10*). These studies revealed that the ABA receptor family comprises small proteins that function in different cell types with different affinities for ABA and, together with ABA, form a ternary complex with a protein phosphatase (*4*). When ABA or its synthetic analogs bind to the receptor, a portion of the protein acts as a gate and swings down and blocks the active site of the phosphatase enzyme.

The study of Vaidya *et al.* combined chemical genomic screens and synthetic organic chemistry guided by protein crystal structures to identify a new generation of ABA receptor agonists that are active at lower concentrations in the two major phylogenetic categories of land plants: dicots (such as soybean) and monocots (such as wheat). Through a salt bridge to a specific lysine in the ABA receptor, the new compounds add a "lock" to the latch on the gate, keeping the phosphatase inhibited. This leads to an increase in the phosphorylation state of a specific protein kinase that initiates a signaling pathway that leads to drought tolerance.

The ABA receptor gene family has 14 members, each of which is expressed at different amounts in different plant tissues and has different affinities for ABA. Therefore, because of this functional redundancy, finding compounds that bind to each of the receptors with different specificities and affinity is an important but challenging goal for basic research purposes as well as for future applications in lowering water use in agriculture.

Even though most plants die when their water content drops below ~75%, all plant cells encode in their DNA the proteins that allow their cells to stay alive at 12% water content, because plants desiccate embryos to form seeds as a natural and carefully programmed part of their life cycle. However, staying alive and producing high yields of food even during periods of drought are two different outcomes. ABA receptor activation controls water dynamics in all cells of the plant, but the loss of water through transpiration via leaf guard cells during the noonday sun while carbon dioxide is taken up through these same pores creates a paradoxical competing situation of maximizing crop yields (by promoting photosynthesis) while minimizing water needs. Perhaps an armada of technologies that chemically and genetically alter the various ABA receptors throughout the plant can hasten a "blue" revolution in water use efficiency that could be as important for our future as the green revolution that arose from the scientific research of decades past. ■

[1]Departments of BioSciences and Chemistry, Rice University, Houston, TX, USA. [2]Department of Biochemistry, University of Wisconsin, Madison, WI, USA. Email: georgep@rice.edu, msussman@wisc.edu

REFERENCES AND NOTES

1. X. Wu, L. K. Dodgen, J. L. Conkle, J. Gan, *Sci. Total Environ.* **536**, 655 (2015).
2. A. S. Vaidya *et al.*, *Science* **366**, 446 (2019).
3. R. E. Hirsch, B. D. Lewis, E. P. Spalding, M. R. Sussman, *Science* **280**, 918 (1998).
4. M. R. Sussman, G. N. Phillips Jr., *Science* **326**, 1356 (2009).
5. H. Fujii *et al.*, *Nature* **462**, 660 (2009).
6. Y. Ma *et al.*, *Science* **324**, 1064 (2009).
7. K. Melcher *et al.*, *Nature* **462**, 602 (2009).
8. N. Nishimura *et al.*, *Science* **326**, 1373 (2009).
9. S. Y. Park *et al.*, *Science* **324**, 1068 (2009).
10. J. Santiago *et al.*, *Nature* **462**, 665 (2009).

10.1126/science.aaz4857

EVOLUTION

The treacheries of adaptation

As fitness rises during adaptive evolution, the cost of mutation may escalate as well

By **Craig R. Miller**

One of Darwin's great insights was that he took the widespread observation that organisms are exceptionally well-suited to their environment and turned it on its head. He argued that behind the constructive process of adaptation lies, counterintuitively, a destructive one: Progeny with favorable variations obscure the many progeny who are less well suited and either do not survive or, at best, have fewer offspring. This insight—which Darwin owes in part to Thomas Malthus and shares with Alfred Wallace—is a cornerstone in the theory of evolution by natural selection and raises an important question in biology: What is the nature of the trade-off between the capacity for adaptation and the cost of producing less-fit offspring? On page 490 of this issue, Johnson *et al.* (*1*) find that as adaptation proceeds and fitness gains are expected to diminish, the cost of mutation becomes more severe.

The modern framing of adaptation imagines genotypes as hyperdimensional Cartesian coordinates and fitness as the elevation above them. This gives rise to a fitness landscape that a population must traverse if it is to adapt (see the figure). The landscape metaphor aids visualization of the trade-off. The capacity for adaptive evolution is determined by the curvature of the uphill terrain. This curvature results from how mutations interact, called epistasis. A wealth of microbial evolution experiments (*2*–*6*) suggest that the uphill curvature is usually concave, giving rise to "diminishing returns epistasis" as fitness increases. Thus, the fitness effect of a beneficial mutation is not the same across backgrounds; it gets smaller as background fitness increases. The cost of mutations among offspring is also related to curvature; if the vicinity around a genotype is concave in all directions, the costs will be small and the genotype is considered robust. High robustness occurs when mutations have little effect on phenotype (*7*). Compared to the uphill terrain, the curvature of the downslope terrain has been less well studied in the context of adaptation.

Johnson *et al.* sought to characterize the landscape around genetically distinct variants by estimating the fitness effect of insertion mutations in yeast (*Saccharomyces cerevisiae*). They found that most mutations were deleterious (that is, they reduced fitness) and had larger deleterious effects on higher-fitness backgrounds. They called this pattern "increasing cost epistasis."

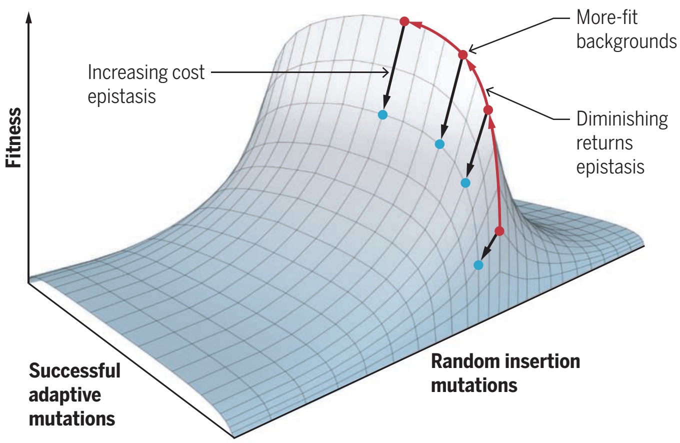
Examining mutations individually, they found that most follow the pattern of increasing cost epistasis. A substantial minority, however, either showed the opposite pattern (diminishing costs) or no trend at all. The same backgrounds were previously the subject of quantitative trait loci (QTL) mapping (*8*), whereby genetic differences at tens of thousands of sites are used to find genome locations associated with measurable (quantitative) traits. This allowed the authors to ask how well a QTL model explains fitness effects and compare it to models that include background fitness or both. They found that the QTL model is best for some mutations, the background fitness model is best for others, and, for most, the best model includes both QTLs and background fitness. That QTLs were not consistently the best explanation is surprising because they were very good at explaining variation in quantitative traits (*8*). The implication is that although there is an overarching pattern of increasing cost epistasis, there is also interesting variation at the level of individual mutations. This also suggests that models with mechanistic underpinnings (*6*, *9*) may be more successful than simple models of epistasis (*10*, *11*) in explaining adaptive evolution.

The results of Johnson *et al.* are unexpected because they imply different curvatures of the fitness landscape depending on what type of mutations are tested: Random insertion mutations follow a convex function on fitness, whereas successful beneficial mutations during adaptation follow a concave one (see the figure). Although epistasis scales with fitness in both cases, the relationship is in opposite directions. How can this be? The authors argue that increasing cost epistasis may arise from metabolic flux, whereby a deleterious mutation in a sequential pathway has lesser negative consequences when other enzymes in the pathway have already been adversely affected. If low-fitness backgrounds have less-functional metabolic pathways, disruptions through mutation can do less damage. But this argument, when reversed, is problematic. It implies that beneficial mutations on higher-flux pathways should generate synergies (or convex surfaces)—such mutations are observed (*2*), but rarely. Part of the explanation may be that some adaptive mutations in experimental evolution succeed by disrupting expendable pathways; once such pathways are disrupted, further mutations will be of diminishing benefit. Indeed, beneficial disruptive mutations are not uncommon (*5*, *12*).

The different curvatures may also reflect how mutations are sampled. Insertion mutations disrupt random pathways; if disrupting a pathway is usually detrimental, it will tend to generate increasing cost epistasis. To quote the singer-songwriter

A fitness landscape with two types of epistasis

Disruptive insertion mutations have greater deleterious effects on more-fit backgrounds, called increasing cost epistasis. By contrast, most adaptation experiments show a pattern of diminishing returns epistasis, whereby mutations confer smaller benefit as fitness increases. The different curvatures may reflect differences in how deleterious and beneficial mutations are sampled.



GRAPHIC: C. BICKEL/*SCIENCE*

Department of Biological Sciences, Institute for Modeling Collaboration & Innovation, University of Idaho, Moscow, ID 83844, USA. Email: crmiller@uidaho.edu

Townes Van Zandt, "you don't need no engine to go downhill." Conversely, adaptation is biased. When disruptive mutations can confer fitness gains, they will be selected—even if they return only diminishing benefit.

Returning to Darwin's insight and the question of trade-offs in adaptation, the study of Johnson *et al.* suggests that as organisms adapt through natural selection of beneficial mutations, they will concurrently suffer an escalating burden of producing less-fit offspring. The authors speculate that increasing costs, paired with the reality of diminishing gains, may arrest adaptation before a fitness optimum is reached. An important test of the authors' claim will be to assess if increasing cost epistasis is also observed during an adaptive walk, where the genotypes of increasing fitness differ by a far smaller number of mutations. Moreover, it is unclear how often mutation rates will be high enough and beneficial effects small enough

"...as adaptation proceeds and fitness gains are expected to diminish, the cost of mutation becomes more severe."

for the forces to counterbalance. There is also accumulating evidence that some biological features—gene expression patterns and protein stabilization by chaperones, for example—are robust, meaning they have apparently escaped the treacheries of increasing cost epistasis (*7*). This raises the question, how have they done so? ■

REFERENCES AND NOTES

1. M. S. Johnson *et al.*, *Science* **366**, 490 (2019).
2. A. I. Khan, D. M. Dinh, D. Schneider, R. E. Lenski, T. F. Cooper, *Science* **332**, 1193 (2011).
3. H. H. Chou, H.-C. Chiu, N. F. Delaney, D. Segrè, C. J. Marx, *Science* **332**, 1190 (2011).
4. D. R. Rokyta *et al.*, *PLOS Genet.* **7**, e1002075 (2011).
5. S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, *Science* **344**, 1519 (2014).
6. X. Wei, J. Zhang, *Mol. Biol. Evol.* **36**, 1008 (2019).
7. J. L. Payne, A. Wagner, *Nat. Rev. Genet.* **20**, 24 (2019).
8. J. S. Bloom, I. M. Ehrenreich, W. T. Loo, T.-L. V. Lite, L. Kruglyak, *Nature* **494**, 234 (2013).
9. H.-C. Chiu, C. J. Marx, D. Segrè, *Proc. R. Soc. London Ser. B* **279**, 4156 (2012).
10. P. Joyce, D. R. Rokyta, C. J. Beisel, H. A. Orr, *Genetics* **180**, 1627 (2008).
11. C. R. Miller, J. T. Van Leuven, H. A. Wichman, P. Joyce, *Theor. Popul. Biol.* **122**, 97 (2018).
12. O. Tenaillon *et al.*, *Science* **335**, 457 (2012).

ACKNOWLEDGMENTS

Thanks to C. Marx and J. Bull for help with writing. Support comes from NIH grants R01 GM076040 and P20 GM104420, and NSF grant OIA-1736253.

10.1126/science.aaz5189

IMMUNOLOGY

Immune cells for microbiota surveillance

Commensal microbes regulate specialized T cells, which promote wound healing in skin

By **Julia Oh** *and* **Derya Unutmaz**

The immune system has coevolved with the microbial community that inhabits body surfaces and mucosal barriers. Although this commensal microbiota is critical for maintaining healthy host physiology, it can cause pathology when the body surface barriers are breached. How the immune system maintains this homeostasis with microbiota remains poorly understood. Specialized immune cells, called mucosal-associated invariant T (MAIT) cells, specifically recognize and respond to microbial metabolites and are thought to be important in microbial defense, although their function remains unclear. On pages 445 and 494 of this issue, Constantinides *et al.* (*1*) and Legoux *et al.* (*2*), respectively, show that commensal bacteria control development of MAIT cells in the thymus and their expansion within mucosal tissues. The development of MAIT cells depends on a specific developmental window of early-life exposure to defined microbial communities, and a distinct MAIT cell subset in the skin promotes wound healing.

Conventional T cells recognize peptide antigens presented by major histocompatibility complex (MHC) molecules. However, MAIT cells are nonclassical T cells because they are stimulated by nonpeptide antigens—specifically, vitamin B2 precursor derivatives produced by many bacteria that are bound to an MHC-like protein called MR1. Several studies suggest that MAIT cells play an important role in immunity for controlling bacterial, fungal, and viral infections (*3*). MAIT cells have been further implicated in nonmicrobial diseases, including autoimmune diseases, and have potential roles in tumor immunity.

The intriguing paucity of MAIT cells in germ-free (GF) mice, which are reared in microbe-free conditions, suggested that they require an established microbiota to develop (*4*). Constantinides *et al.* and Legoux *et al.* provide extensive evidence that the development of MAIT cells is tightly linked to the availability of the commensal microbial derivatives, which enable both their development in the thymus and further tissue-specific expansion. In mice, the type and timing of microbial colonization also determine the MAIT cell frequencies in tissues. These findings provide further clues about the high variation in the frequency of MAIT cells in humans, which can vary by 40-fold (*5*).

Jackson Laboratory for Genomic Medicine and University of Connecticut School of Medicine, Farmington, CT 06032, USA. Email: derya@mac.com

The findings by Constantinides *et al.* and Legoux *et al.* emphasize the role of microbial colonization in early life because they identify a specific developmental window after birth for MAIT cell education in mice (see the figure). In humans, MAIT cells begin to develop in utero, and all newborns are rapidly colonized by a diverse set of commensal bacterial species and strains (*6*). Although many bacterial species are riboflavin-synthesizing, the synthesis of intermediate metabolites that activate MAIT cells may vary in different bacterial species (*7*). Constantinides *et al.* also found that certain bacteria species—such as Enterobacteriaceae, including *Proteus* and *Klebsiella* species—were most efficient in MAIT cell development, although such interactions appear to have local constraints. For example, *P. mirabilis* colonization of both GF neonates and adults was sufficient to induce mature MAIT cells in the thymus but not in tissues. However, it is not yet known how in humans the dynamic nature of the microbiota through a lifetime contributes to the equally dynamic MAIT cell maturation and expansion, which is gradual during early childhood, peaking in early adulthood followed by decline in the elderly (*8*). Thus, in humans, the frequency and tissue localization of MAIT cells at different stages in life are likely determined both by early life imprinting and expansion based on the presence of different bacterial species colonizing the barrier sites.

A key remaining question is how MAIT cells can respond to a diverse set of microbiota and yet discriminate pathogenic microorganisms from commensals. It appears that MAIT cells have acquired diverse functional programs and can tune their outputs by integrating both antigen

and nonantigen signals from the environment. They can produce cytokines that are associated with T cell subsets such as T helper 1 (T_H1) and T_H17 cells that produce proinflammatory cytokines and have potent cytotoxic activity toward cells infected with bacteria (*9*). Constantinides *et al.* extend this functional diversity by demonstrating that MAIT cell subsets resident in the skin are endowed with a transcriptional program of wound healing or tissue repair, which is distinct from MAIT cells in other tissues. The surveillance of skin commensal–derived metabolites and the presence of a cytokine milieu at a skin wound or damage possibly trigger MAIT cells to induce their tissue repair functions. It is not yet clear whether these transcriptional programs are established after migration to the skin or whether MAIT cells are preprogrammed.

> **"These findings are also relevant to determining the precise role of MAIT cells in... chronic diseases."**

On the basis of expression of co-receptors, MAIT cells can either express CD4 or CD8 or can be $CD4^-CD8^-$ [double negative (DN)] (*9*). Although the role of CD4 and CD8 expression in MAIT cell biology is unclear, it is curious that MAIT cells in mouse skin and other barrier sites, such as the lungs and gut, are almost exclusively DN. Human skin also contains $CD8^+$ MAIT cells (*10*); thus, it remains unclear to what extent DN MAIT cells are a separate subset with alternative functions and whether $CD8^+$ MAIT cells in human skin are also involved in wound repair.

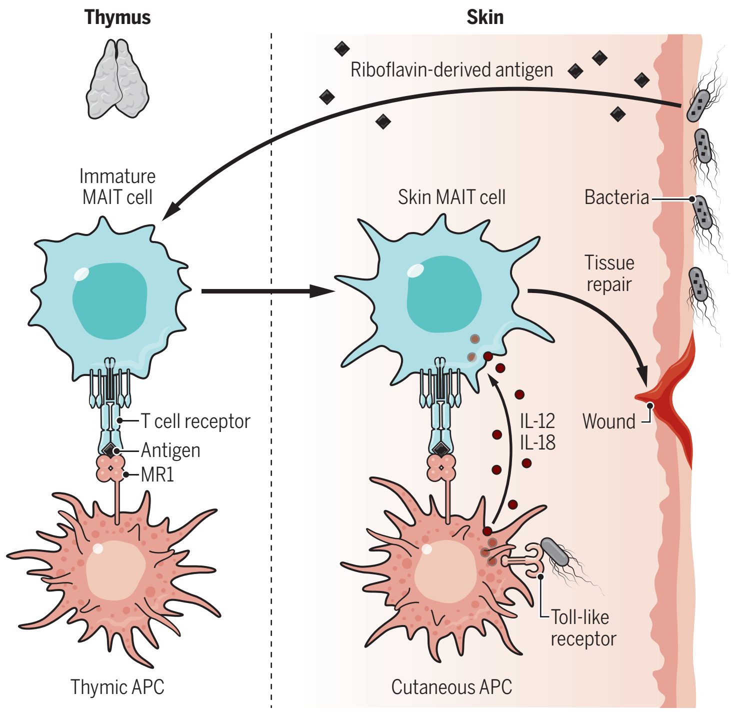
Commensal skin microbes, such as *Staphylococcus epidermidis,* have important roles in establishing immune tolerance in mice, tuning resident T lymphocyte function and inflammatory responses, and conferring protective immunity against skin pathogens (*11*, *12*). They may also play a role in triggering MAIT cells for local wound healing. It is important to note, however, that the human skin microbiota is both temporally and spatially dynamic within an individual over a life span, largely driven initially by maternal inheritance of skin microbes (*13*) and subsequently by physiologic maturation and characteristics of the skin (*14*). Microbes such as *S. epidermidis* can vary in abundance and genetic diversity at different skin sites. Moreover, some skin microbes such as *Cutibacterium acnes* and many Firmicutes have low MAIT cell activation capacity in vitro or lack the classical riboflavin synthesis pathway (*7*). Whether such differences in the abundance or composition of skin-resident microbiota correspondingly affect MAIT cell activation in skin and alter their role in skin wound healing and tissue repair will be a fascinating sequel.

Many questions remain about how MAIT cells are optimally activated to elicit their regional programs in different tissues. Moreover, the degree of plasticity between different functional states will be interesting to understand. An important characteristic of MAIT cells is the expression of the interleukin-18 receptor (IL-18R) during development, which allows their stimulation by IL-18, a proinflammatory cytokine induced from innate cells during infections (*4*, *9*). Cutaneous MAIT cells also express IL-18R, and neutralizing IL-18 reduces the MAIT cell response to *S. epidermidis*, indicating that although MAIT cell response is strictly antigen stimulation–dependent, IL-18 cosignaling was also required for their optimal expansion (*1*). Given that MAIT cells in the skin and mucosal tissues may be in constant proximity to microbial antigens, the question becomes, why are they not chronically activated? The answer may lie in understanding both environmental and genetic regulation of the production of stimulatory metabolites by microbes and other environmental signals from cytokines or other bacterial products such as Toll-like receptor ligands, either directly or through the stimulation of antigen-presenting cells.

The striking observations from Constantinides *et al.* and Legoux *et al.* that the microbiota serve as an important regulator of MAIT cell thymic development, abundance, and tissue specificity provide a framework to understand their tremendous dynamic range between individuals during life. These findings are also relevant to determining the precise role of MAIT cells in infections, tumor immunity, and chronic diseases. It is in some ways ingenious that the immune system uses microbiota to decide the quality and quantity of MAIT cells, which in turn determine the surveillance and homeostasis of the microbial ecosystem. This personalized MAIT cell frequency and function is disrupted during aging, certain chronic infections such as HIV, and other chronic diseases. Whether manipulating the microbiota can reconstitute such personalized functionality warrants further investigation. ■

Microbiota regulate MAIT cell development and function

Mucosal-associated invariant T (MAIT) cell development in the thymus and migration to the skin are initiated by bacteria-derived riboflavin antigens. Resident cutaneous MAIT cells are activated by a combination of bacteria-derived antigens displayed by MR1-expressing antigen-presenting cells (APCs) and cytokines [such as interleukin-12 (IL-12) and IL-18] that are released locally. Activated MAIT cells then release factors that promote skin repair. MR1, MHC-related protein 1.

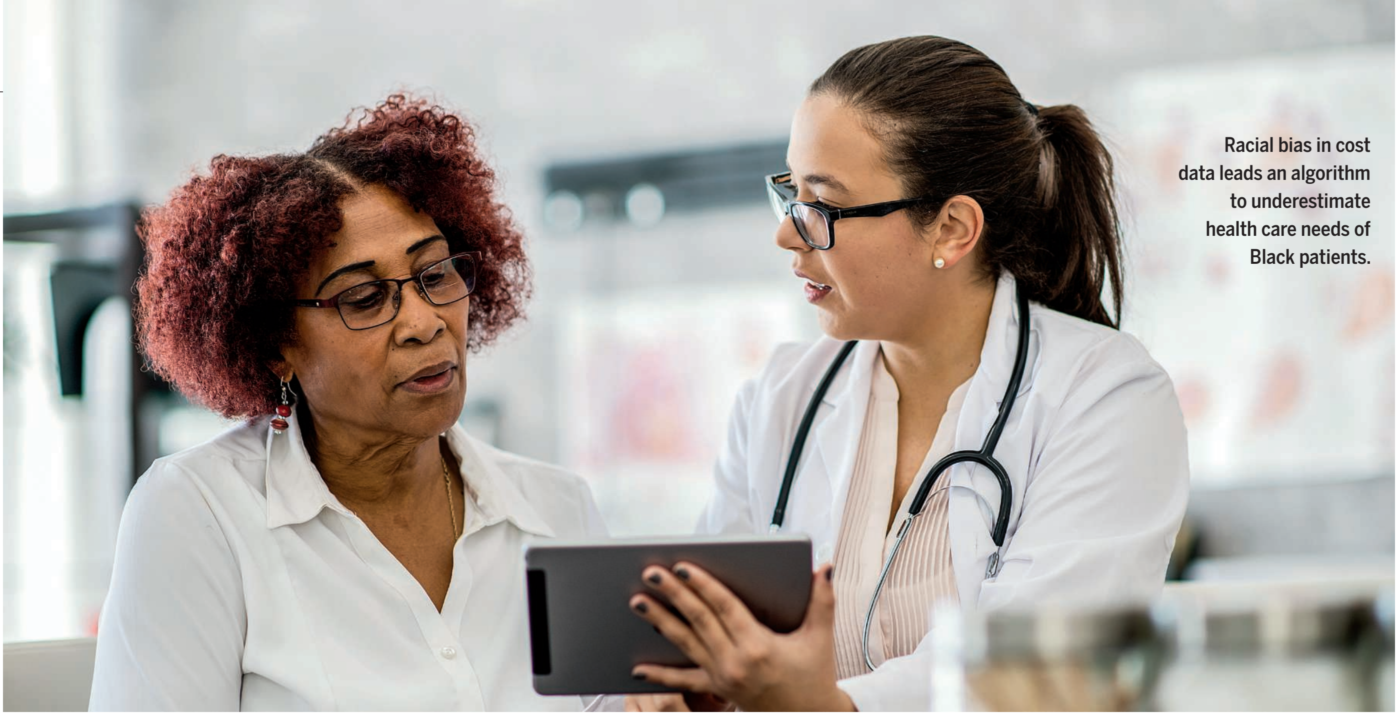


GRAPHIC: A. KITTERMAN/*SCIENCE*

REFERENCES AND NOTES

1. M. G. Constantinides *et al.*, *Science* **366**, 445 (2019).
2. F. Legoux *et al.*, *Science*, **366**, 494 (2019).
3. M. Salou *et al.*, *Curr. Opin. Immunol.* **48**, 7 (2017).
4. H. F. Koay *et al.*, *Nat. Immunol.* **17**, 1300 (2016).
5. G. Ben Youssef *et al.*, *J. Exp. Med.* **215**, 459 (2018).
6. A. K. Simon *et al.*, *Proc. Biol. Sci.* **282**, 20143085 (2015).
7. C. Tastan *et al.*, *Mucosal Immunol.* **11**, 1591 (2018).
8. L. J. Walker *et al.*, *Scand. J. Immunol.* **80**, 462 (2014).
9. D. I. Godfrey *et al.*, *Nat. Immunol.* **20**, 1110 (2019).
10. M. B. M. Teunissen *et al.*, *J. Invest. Dermatol.* **134**, 2898 (2014).
11. S. Naik *et al.*, *Science* **337**, 1115 (2012).
12. A. Nosbaum *et al.*, *J. Immunol.* **196**, 2010 (2016).
13. G. Dominguez-Bello *et al.*, *Nat. Med.* **22**, 250 (2016).
14. A. L. Byrd *et al.*, *Nat. Rev. Microbiol.* **16**, 143 (2018).

10.1126/science aaz4014



Racial bias in cost data leads an algorithm to underestimate health care needs of Black patients.

SOCIAL SCIENCE

Assessing risk, automating racism

A health care algorithm reflects underlying racial bias in society

By **Ruha Benjamin**

As more organizations and industries adopt digital tools to identify risk and allocate resources, the automation of racial discrimination is a growing concern. Social scientists have been at the forefront of studying the historical, political, economic, and ethical dimensions of such tools (*1–3*). But most analysts do not have access to widely used proprietary algorithms and so cannot typically identify the precise mechanisms that produce disparate outcomes. On page 447 of this issue, Obermeyer *et al.* (*4*) report one of the first studies to examine the outputs and inputs of an algorithm that predicts health risk, and influences treatment, of millions of people. They found that because the tool was designed to predict the cost of care as a proxy for health needs, Black patients with the same risk score as White patients tend to be much sicker, because providers spend much less on their care overall. This study contributes greatly to a more socially conscious approach to technology development, demonstrating how a seemingly benign choice of label (that is, health cost) initiates a process with potentially life-threatening results. Whereas in a previous era, the intention to deepen racial inequities was more explicit, today coded inequity is perpetuated precisely because those who design and adopt such tools are not thinking carefully about systemic racism.

Obermeyer *et al.* gained access to the training data, algorithm, and contextual data for one of the largest commercial tools used by health insurers to assess the health profiles for millions of patients. The purpose of the tool is to identify a subset of patients who require additional attention for complex health needs before the situation becomes too dire and costly. Given increased pressure by the Affordable Care Act to minimize spending, most hospital systems now utilize predictive tools to decide how to invest resources. In addition to identifying the precise mechanism that produces biased predictions, Obermeyer *et al.* were able to quantify the racial disparity and create alternative algorithmic predictors.

Practically speaking, their finding means that if two people have the same risk score that indicates they do not need to be enrolled in a "high-risk management program," the health of the Black patient is likely much worse than that of their White counterpart. According to Obermeyer *et al.*, if the predictive tool were recalibrated to actual needs on the basis of the number and severity of active chronic illnesses, then twice as many Black patients would be identified for intervention. Notably, the researchers went well beyond the algorithm developers by constructing a more fine-grained measure of health outcomes, by extracting and cleaning data from electronic health records to determine the severity, not just the number, of conditions. Crucially, they found that so long as the tool remains effective at predicting costs, the outputs will continue to be racially biased by design, even as they may not explicitly attempt to take race into account. For this reason, Obermeyer *et al.* engage the literature on "problem formulation," which illustrates that depending on how one defines the problem to be solved—whether to lower health care costs or to increase access to care—the outcomes will vary considerably.

To grasp the broader implications of the study, consider this hypothetical: The year is 1951 and an African American mother of five, Henrietta Lacks, goes to Johns Hopkins Hospital with pain, bleeding, and a knot in her stomach. After Lacks is tested and treated with radium tubes, she is "digitally triaged" (*2*) using a new state-of-the-art risk assessment tool that suggests to hospital staff the next course of action. Because the tool assesses risk using the predicted cost of care, and because far less has commonly been spent on Black patients despite their actual needs, the automated system underestimates the level of attention Lacks needs. On the basis of the results, she is discharged, her health rapidly deteriorates,

Department of African American Studies, Princeton University, Princeton, NJ, USA. Email: ruha@princeton.edu

and, by the time she returns, the cancer has advanced considerably, and she dies.

This fictional scenario ends in much the same way as it did in reality, as those familiar with Lacks's story know well (*5–7*). But rather than getting assessed by a seemingly race-neutral algorithm applied to all patients in a colorblind manner, she was admitted into the Negro wing of Johns Hopkins Hospital during a time when explicit forms of racial discrimination were sanctioned by law and custom—a system commonly known as Jim Crow. However, these are not two distinct processes, but rather Jim Crow practices feed the "New Jim Code"—automated systems that hide, speed, and deepen racial discrimination behind a veneer of technical neutrality (*1*).

Data used to train automated systems are typically historic and, in the context of health care, this history entails segregated hospital facilities, racist medical curricula, and unequal insurance structures, among other factors. Yet many industries and organizations well beyond health care are incorporating automated tools, from education and banking to policing and housing, with the promise that algorithmic decisions are less biased than their human counterpart. But human decisions comprise the data and shape the design of algorithms, now hidden by the promise of neutrality and with the power to unjustly discriminate at a much larger scale than biased individuals.

For example, although the Fair Housing Act of 1968 sought to protect people from discrimination when they rent or buy a home, today social media platforms allow marketers to explicitly target advertisements by race, excluding racialized groups from the housing market without penalty (*8*). Although the federal government brought a suit against Facebook for facilitating digital discrimination in this manner, more recently the U.S. Department of Housing and Urban Development introduced a rule that would make it harder to fight algorithmic discrimination by lenders, landlords, and others in the housing industry. And unlike the algorithm studied by Obermeyer *et al.*, which used a proxy for race that produced a racial disparity, targeted ads allow for explicit racial exclusion, which violates Facebook's own policies. Yet investigators found that the company continued approving ads excluding "African Americans, mothers of high school kids, people interested in wheelchair ramps, Jews, expats from Argentina and Spanish speakers," all within minutes of an ad submission (*8*). So, whether it is a federal law or a company policy, top-down reform does not by itself dampen discrimination.

Labels matter greatly, not only in algorithm design but also in algorithm analysis. Black patients do not "cost less," so much as they are valued less (*9*). It is not "something about the interactions that Black patients have with the healthcare system" that leads to poor care, but the persistence of structural and interpersonal racism. Even health care providers hold racist ideas, which are passed down to medical students despite an oath to "do no harm" (*10*). The trope of the "noncompliant (Black) patient" is yet another way that hospital staff stigmatize those who have reason to question medical authority (*11*, *12*). But a "lack of trust" on the part of Black patients is not the issue; instead, it is a lack of trustworthiness on the part of the medical industry (*13*). The very designation "Tuskegee study" rather than the official name, U.S. Public Health Service Syphilis Study at Tuskegee, continues to hide the agents of harm. Obermeyer *et al.* mention some of this context, but passive and sanitized descriptions continue to hide the very social processes that make their study consequential. Labels matter.

As researchers build on this analysis, it is important that the "bias" of algorithms does not overshadow the discriminatory context that makes automated tools so important in the first place. If individuals and institutions valued Black people more, they would not "cost less," and thus this tool might work similarly for all. Beyond this case, it is vital to develop tools that move from assessing individual risk to evaluating the production of risk by institutions so that, ultimately, the public can hold them accountable for harmful outcomes. ■

REFERENCES AND NOTES

1. R. Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code* (Polity Press, 2019).
2. V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin's Press, 2018).
3. S. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (NYU Press, 2018).
4. Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, *Science* **366**, 447 (2019).
5. K. Holloway, *Private Bodies, Public Texts: Race, Gender, and a Cultural Bioethics* (Duke Univ. Press, 2011).
6. H. Landecker, *Sci. Context* **12**, 203 (1999).
7. R. Skloot, *The Immortal Life of Henrietta Lacks* (Broadway Books, 2011).
8. J. Angwin, A. Tobin, M. Varner, "Facebook (still) letting housing advertisers exclude users by race," *ProPublica*, 21 November 2017; www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.
9. E. Glaude Jr., *Democracy in Black: How Race Still Enslaves the American Soul* (Crown Publishers, 2016).
10. K. M. Bridges, *Reproducing Race: An Ethnography of Pregnancy as a Site of Racialization* (Univ. California Press, 2011).
11. A. Nelson, *Body and Soul: The Black Panther Party and the Fight Against Medical Discrimination* (Univ. Minnesota Press, 2011).
12. H. Washington, *Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present* (Harlem Moon, Broadway Books, 2006).
13. R. Benjamin, *People's Science: Bodies and Rights on the Stem Cell Frontier* (Stanford Univ. Press, 2013).

10.1126/science.aaz3873

BATTERY TECHNOLOGY

The coming electric vehicle transformation

A future electric transportation market will depend on battery innovation

By **George Crabtree**[1,2]

Electric vehicles are poised to transform nearly every aspect of transportation, including fuel, carbon emissions, costs, repairs, and driving habits. The primary impetus now is decarbonization to address the climate change emergency, but it soon may shift to economics because electric vehicles are anticipated to be cheaper and higher-performing than gasoline cars. The questions are not if, but how far, electrification will go. What will its impact be on the energy system and on geoeconomics? What are the challenges of developing better batteries and securing the materials supply chain to support new battery technology?

The signs of vehicle electrification are growing. By 2025, Norway aims to have 100% of its cars be either an electric or plug-in hybrid unit, and the Netherlands plans to ban all gasoline and diesel car sales by the same year. By 2030, Germany plans to ban internal combustion engines, and by 2040, France and Great Britain aim to end their gasoline and diesel car sales. The most aggressive electric vehicle targets are those set by China, which has almost half the global electric vehicle stock and where 1.1 million electric vehicles were sold in 2018. Europe and the United States each have just over 20% of the global stock, with electric car sales of 380,000 and 375,000 units, respectively, in 2018 (*1*, *2*).

How far electrification will go depends primarily on a single factor—battery technology. In comparing electric with gasoline vehicles, all the downsides for electric arise from the battery. Purchase price, range, charging time, lifetime, and safety are all battery-driven handicaps. On the upside, electric vehicles have lower greenhouse gas emissions, provided the electricity grid that supports them is powered by renewable energy [the renewable share of global electricity is up from 22% in 2001 to 33% today (*3*), with Europe at 36%, China at 26%, and

An electric car in Milan, Italy, gets a charge. Grid-connected renewable energy systems, improved energy storage, and new battery technology will accelerate the electrification of transportation.

the United States at 18% (*4*)]. Moreover, the operation and maintenance costs of electric vehicles are substantially lower than for gasoline cars. Today, for high-mileage cars such as taxis, which typically travel 70,000 miles/year, the total cost of ownership of an electric vehicle, including purchase price, insurance, fuel, and maintenance, is much lower than for a gasoline car. This means that government and commercial fleets used for local service likely will convert to electric to save money, a major step in the electrification trajectory. To reach cost parity with personal gasoline cars, which typically travel 12,000 to 15,000 miles/year, battery prices must decline to near \$100/kWh from the present value of \$180 to \$200/kWh. Projections of the year of cost parity for electric vehicles with gasoline cars globally range from 2022 to 2026 (*5*, *6*). At that point, economics could well take over as the primary impetus for electrification, and electric vehicles would then be on a path to transportation dominance.

[1]Joint Center for Energy Storage Research, Argonne National Laboratory, Lemont, IL, USA. [2]University of Illinois at Chicago, Chicago IL, USA. Email: crabtree@anl.gov

IMPACT ON ENERGY SYSTEM

Electric vehicles will need to be charged from the grid, which may create as much as a 20 to 38% increase in electricity demand by 2050 (*7*). In developed countries, this should provide revenue for utilities to accelerate transformation to a grid-connected renewable energy system with extensive energy storage and to digital energy management. In developing countries, the increased electricity demand could spur the first-time installation of modern grids that are unencumbered by the legacy of the older, less functional grids of the developed world. Beyond electricity, electric vehicles require a massive rollout of charging stations, which could stimulate local economic and job growth.

Electric vehicles also should bring a welcome flexibility to the energy system. Untied from oil and gasoline, they would run on whatever powers the grid—sunlight, wind, natural gas, nuclear power, or hydropower. This removes a fundamental dependence of transportation on oil, including substantial amounts of foreign oil in many countries. Electricity is fundamentally a local product, not amenable to long-distance trade, so domestic economies should reap the economic and job benefits now held by foreign oil interests. The unification of transportation with electricity creates new horizons of opportunity for the grid as well. Electric vehicles are a readily available distributed energy resource of at least 1000 GWh, which represents 10% of the battery capacity of 100 million vehicles, each with a 100-kWh battery. The potential of this distributed energy resource for demand response and for grid storage has not yet been seriously explored.

IMPACT ON GEOECONOMICS

The electrification of transportation is a watershed moment in energy economics. For more than a century, oil has been the lifeblood of transportation, and the oil industry has grown steadily as transportation has expanded with industrialization and rising standards of living. But oil is abundant in relatively few countries, and these countries assume outsized geoeconomic importance because oil for transportation is a critical societal need. By contrast, sunlight and wind are available everywhere, and electric-

ity generation is mostly a domestic enterprise. The electrification of transportation means that oil will lose one of its critical markets—and with it some of its international economic and political power.

What will replace oil as the lifeblood of transportation? The electrification of transportation creates a new commodity—not electricity, which is already established and abundant around the world, but battery technology. The battery is the key to electric transportation, the focal point for progress, and the open opportunity to determine the future of electric vehicles. Battery innovation is needed to achieve lower purchase price, faster charging, longer range, extended lifetime, and greater safety. These challenges do not yet have obvious solutions, but those who discover them will have substantial power in the battery marketplace.

BATTERY DEVELOPMENT

One of the most promising and disruptive battery innovations is the combination of lithium metal anodes and solid-state electrolytes. Every atom of a lithium metal anode can store and release energy during the charge-discharge cycle, whereas in graphite anodes now used in lithium-ion batteries, only 14% of the atoms (one lithium for every six carbons) can store or release energy. The greater capacity of the lithium metal anode could approximately double the energy density of the lithium-ion battery, extending the driving range of electric vehicles to compete with gasoline cars.

Solid-state electrolytes bring several advantages to lithium-ion batteries (*8*). They are not flammable, eliminating the primary safety hazard of lithium-ion batteries—the thermal runaway reaction that causes batteries to burst into flames if their temperature exceeds about 150°C. Some solid-state electrolytes, including sulfides such as Li_2S–P_2S_5 (LPS) and garnets such as $Li_7La_3Zr_2O_{12}$ (LLZO), have high lithium-ion conductivity at room temperature, enabling the high-power performance needed for fast charging. Solid-state electrolytes conduct heat better than liquid electrolytes, protecting against the development of "hot spots" that trigger degradation and shorten battery life. In addition, the mechanical rigidity of solid-state electrolytes can block the growth of dendrites that form on the lithium metal anode surface and grow across liquid electrolytes to the cathode, shorting out the battery. These benefits of solid-state electrolytes are balanced by still-unresolved research challenges, including narrow working voltage windows, high reactivity with lithium anodes, and long-term stability.

There is now an intense drive to develop lithium metal anodes and solid-state electrolytes spanning academic, government, and industrial laboratories. Toyota announced its intention to have batteries with lithium anodes and solid-state electrolytes ready for electric vehicles by the early 2020s (*9*). The combination of lithium metal anodes with solid-state electrolytes would mark the first disruptive step in lithium-ion battery development, breaking a three-decade pattern of steady incremental advances in performance and cost (*10*).

MATERIAL SUPPLY CHAINS

Lithium, cobalt, manganese, nickel, and graphite are essential for battery technology, and some of these elements are found in only a few places in the world, not unlike oil (*11*, *12*). The expected rapid increase in electric vehicle sales could threaten the supply chains for lithium, cobalt, and graphite in the short term because of the time required to ramp up new materials production and the relative scarcity of geographic sources. In the longer term, there are adequate resources in Earth's crust if lithium-ion batteries are recycled. Currently, less than 5% of Li-ion batteries are recycled, compared to more than 99.5% of lead-acid batteries. (*13*) Research and development to develop Li-ion battery recycling technology is an urgent need.

Batteries and their supply chains are the new oil; leadership in the battery and electric vehicle market requires strategically securing not only battery technology but also the battery materials supply chain. Recycling can play a substantial role in securing the supply chain for lithium-ion batteries, lowering costs by as much as 20% and supplying as much as 50% of the required materials (*12*). The nation or region that leads battery technology and secures its supply chain will have outsized influence on geoeconomics and world development.

GLOBAL LANDSCAPE

Europe has grasped the electric vehicle opportunity, driven by its strict carbon emission requirements for future vehicles. The United States, by contrast, has proposed weakening its carbon emission requirements, and target dates for electrification of transportation are correspondingly farther out. In the International Energy Agency's New Policy Scenario (*1*), electric vehicles are projected to reach 26% of new car sales in Europe by 2030, but only 8% in the United States. China slightly leads Europe, with a 28% share of electric vehicles in 2030. In addition, China has moved strategically to secure its battery supply chain (*11*, *12*). China now has the largest electrical vehicle market and the largest battery manufacturing enterprise in the world, amounting to 60% of the global capacity (*14*). It is well positioned to benefit economically and politically from the coming global electrification of transportation.

The electrification of transportation is far from complete. Buses, long-haul trucking, air taxis, and regional flight (*15*) remain relatively untapped opportunities. Batteries still must overcome challenges in cost, range, charging speed, safety, and lifetime for electric vehicles to dominate the market. Recycling is critical to sustainable supply chains but is still in its infancy. There are enormous opportunities for innovation in discovering solutions to these fundamental challenges. The innovating countries and regions will reap enduring economic and geopolitical benefits. ■

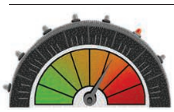
REFERENCES AND NOTES

1. International Energy Agency, Global EV Outlook 2019 (May 2019).
2. A. Steer, How China Raised the Stakes for Electric Vehicles (World Resources Institute, 10 December 2018); www.wri.org/blog/2018/12/how-china-raised-stakes-electric-vehicles.
3. I. R. E. N. A. Renewable Capacity Statistics, 2019, Renewable Capacity Highlights (31 March 2019); www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Mar/RE_capacity_highlights_2019.pdf?la=en&hash=BA9D38354390B001DC0CC9BE03EEE559C280013F
4. Enerdata, Global Energy Statistical Yearbook 2019; https://yearbook.enerdata.net/renewables/renewable-in-electricity-production-share.html
5. J. Romm, Plummeting battery prices to make electric cars cheaper than gas cars in 3 years (ThinkProgress, 16 April 2019); https://thinkprogress.org/electric-vehicles-cheaper-gasoline-cars-e4c86bd2aebe/.
6. M. Holland, $100/kWh Tesla Battery Cells This Year, $100/kWh Tesla Battery Packs In 2020 (Clean Technica, 9 June 2018); https://cleantechnica.com/2018/06/09/100-kwh-tesla-battery-cells-this-year-100-kwh-tesla-battery-packs-in-2020/.
7. T. Mai *et al*., Electrification Futures Study: Scenarios of Electric Technology Adoption and Power Consumption for the United States (National Renewable Energy Laboratory, Golden, CO, NREL/TP-6A20-71500); www.nrel.gov/docs/fy18osti/71500.pdf.
8. J. Janek, W. Zeier, *Nat. Energy* **1**, 16141 (2016).
9. J. LeSage, Solid State Batteries: The Next Big Thing in Electric Cars (OilPrice, 26 June 2019); https://oilprice.com/Energy/Energy-General/Solid-State-Batteries-The-Next-Big-Thing-In-Electric-Cars.html#.
10. R. Schmuch, R. Wagner, G. Hörpel, T. Placke, M. Winter, *Nat. Energy* **3**, 267 (2018).
11. E. A. Olivetti, G. Ceder, G. G. Gaustad, X. Fu, *Joule* **1**, 229 (2017).
12. A. Mayyas, D. Steward, M. Mann, *Sustain. Materi. Technol.* **17**, e00087 (2018).
13. M. Jacoby, *Chem. Eng. News* **97**, 28 (2019).
14. J. Pyper, China Poised to Dominate EV Battery Manufacturing (GTM2, 2 May 2019); www.greentechmedia.com/squared/electric-avenue/china-poised-to-dominate-ev-battery-manufacturing.
15. Reuters, All the electric flying machines come home to roost at the Paris Airshow (19 June 2019); www.autoblog.com/2019/06/19/paris-air-show-electric-aircraft/.

ACKNOWLEDGMENTS

This work was supported as part of the Joint Center for Energy Storage Research, an Energy Innovation Hub funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences.

10.1126/science.aax0704



TOMORROW'S EARTH
Read more articles online at scim.ag/TomorrowsEarth

ECOLOGY

Earthworms' place on Earth

A new study provides a global view of earthworm ecology

By **Noah Fierer**

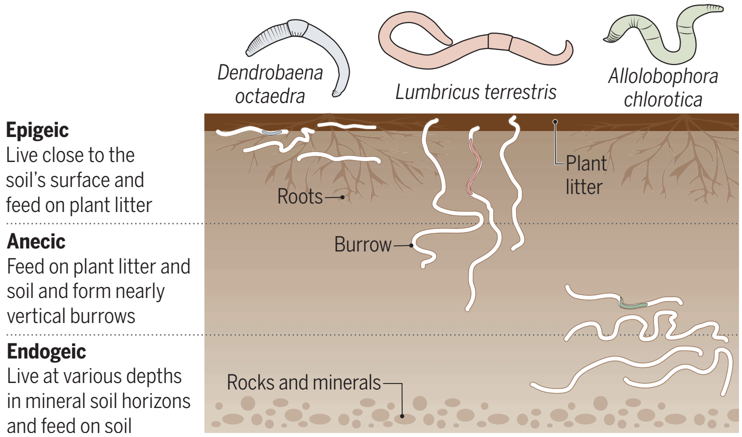
After revolutionizing our understanding of life on Earth, Charles Darwin studied earthworms. In 1881, he published his last scientific book, a treatise on earthworms (*1*) whose sales at the time rivaled those of *On the Origin of Species*. Nearly 140 years later, enthusiasm for earthworms persists, fueled by the recognition of their importance in terrestrial systems as different as backyards and tropical rainforests. On page 480 of this issue, Phillips *et al.* document an impressive group effort by 141 researchers from 35 countries to develop a global-scale atlas of earthworms (*2*). Darwin's legacy continues.

Earthworms, as ecosystem engineers by nature (*3*), influence the structure and functioning of terrestrial ecosystems. Through their burrowing activities, earthworms promote the stabilization of soil particles into aggregates, increase soil porosity, and elevate the rates at which water infiltrates soil during rainfall; reduce erosion of surface soils from hillslopes; and accelerate the movement of gases into or out of soil. Referred to by Aristotle as Earth's intestines, earthworms accelerate organic matter decomposition by ingesting more than 30 times their own weight in soil per day and can rapidly mix large amounts of leaf litter into underlying soil horizons, increasing the release of plant nutrients. The presence of earthworms typically enhances plant growth, including that of most crops, but the magnitude of this effect varies depending on the plant and earthworm species in question (*4*).

A striking example of the impact that earthworms can have on ecosystems comes from studies of temperate and boreal forests that were left devoid of earthworms after the last glacial period. As these earthworm-free ecosystems became colonized by exotic earthworm species, the forests quickly lost the thick litter layers blanketing the soil surface, with corresponding shifts in soil carbon and nitrogen dynamics (*5*). Earthworm invasion can, in turn, cause drastic shifts in plant communities, often leading to the replacement of diverse understory herbaceous and tree seedling communities with lower-diversity plant communities (*6*).

Earthworm ecology

Shown are three main ecological categories of earthworms and examples of resident earthworm species. Not all species fall neatly into these categories, as some earthworms can vary their burrowing and feeding preferences depending on life stage and soil conditions.



GRAPHIC: N. CARY/SCIENCE

Earthworms are found in both managed and unmanaged soils worldwide and can be surprisingly abundant. A single square meter can contain more than 150 individual earthworms with a collective biomass that can exceed 1500 kg per hectare (roughly equivalent to the weight of two adult cattle). Earthworms are also highly diverse, with at least 7000 described species and many more likely awaiting description. Their taxonomic diversity is matched by their morphological and ecological diversity. Earthworms can range in length from just a few centimeters to 3 m (e.g., the endangered *Megascolides australis*) and can be grouped into general categories on the basis of their feeding preferences and burrow architectures, which contribute to species-specific effects on soil properties (see the figure) (*7*).

Despite their clear importance in terrestrial ecosystems, there have been surprisingly few attempts to describe the global biogeographical patterns of earthworms. Unlike many aboveground plant and animal taxa that have been studied at the global scale for decades, ecologists have had only an anecdotal understanding of how the diversity and abundances of earthworms vary on Earth. The Phillips *et al.* study takes an important step toward addressing this knowledge gap.

The authors surveyed ~7000 sites in 56 countries and documented a variety of earthworm distribution patterns. Some were in line with expectations garnered from the way aboveground taxa commonly are distributed on Earth. For example, climatic variables (namely precipitation and temperature) were the best predictors of earthworm diversity and biomass, which is also true for many aboveground taxa. However, other earthworm patterns contrasted with preexisting paradigms in the field of biogeography. For example, unlike many plant and animal taxa (*8*), the diversity of earthworms is not necessarily highest in lower-latitude tropical systems. Southern England might not be considered a hotspot of plant or animal diversity, but it is a veritable earthworm paradise, with soils harboring some of the highest diversity and abundances of earthworms.

Scientific studies are also just as useful for informing researchers of what they don't know. When it comes to earthworms, data show that there is a clear need for more investigations of diversity in particular regions. Although tropical soils appear to have lower earthworm diversity than higher-latitude soils and a greater degree of endemism (i.e., few taxa shared across sites), this pattern might be a product of the paucity of earthworm biologists working in tropical regions. In addition, although the work by Phillips *et al.* demonstrates that earthworm distributions are highly sensitive to climate, it remains unclear how earthworms in soils across the globe will respond to ongoing climate change and what such responses might mean for the functioning of terrestrial ecosystems. Likewise, the local-scale impacts of land-use changes (including pesticide use, tilling, and other agricultural practices) were not the focus of the new study, although ecologists know that such disturbances can

Department of Ecology and Evolutionary Biology, Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80305, USA. Email: noah.fierer@colorado.edu

profoundly alter earthworm populations (*9*). Given the role of earthworms as ecosystem engineers, changes in their diversity and distributions clearly warrant further investigation. Phillips *et al.* provide a roadmap to guide such investigations.

The new study also adds to ongoing efforts to document the diversity of other soil organisms, including bacteria (*10*), fungi (*11*), and nematodes (*12*). These studies are fueled, in part, by the increasing recognition that soil biodiversity is of under-recognized importance to human and ecosystem health. For these reasons, the Global Soil Biodiversity Initiative (*13*) is leading international efforts to gather the scientific knowledge needed to guide and promote the conservation of soil biodiversity. Because soil communities are not necessarily mirror images of aboveground communities, conservation policies designed to protect the more visible aboveground plants and animals might fail to adequately protect the biological diversity hidden beneath our feet. To know how to maximize crop production in agroecosystems, sequester more carbon in soil, or to understand how terrestrial ecosystems will shift in response to human activities, ecologists must learn which soil organisms live where, and what functions they serve in their respective ecosystems. Otherwise, we might not recognize what has been lost before it is too late. ■

"...to understand how terrestrial ecosystems will shift in response to human activities, ecologists must learn which soil organisms live where, and what functions they serve..."

REFERENCES AND NOTES

1. C. Darwin, *The Formation of Vegetable Mould Through the Action of Worms with Some Observations on Their Habits* (John Murray, London, 1881).
2. H. R. P. Phillips *et al.*, *Science* **366**, 480 (2019).
3. C. G. Jones, J. H. Lawton, M. Shachak, *Oikos* **69**, 373 (1994).
4. J. W. van Groenigen *et al.*, *Sci. Rep.* **4**, 6365 (2014).
5. J. M. Crumsey *et al.*, *Biogeochemistry* **126**, 379 (2015).
6. L. E. Frelich *et al.*, *Biol. Invasions* **8**, 1235 (2006).
7. M. Blouin *et al.*, *Eur. J. Soil Sci.* **64**, 161 (2013).
8. K. J. Gaston, *Nature* **405**, 220 (2000).
9. P. J. Bohlen, in *Encyclopedia of Soil Science* (Taylor & Francis, ed. 3, 2017), pp. 701–705.
10. M. Delgado-Baquerizo *et al.*, *Science* **359**, 320 (2018).
11. L. Tedersoo *et al.*, *Science* **346**, 1256688 (2014).
12. J. van den Hoogen *et al.*, *Nature* **572**, 194 (2019).
13. Global Soil Biodiversity Initiative, www.globalsoilbiodiversity.org.

10.1126/science.aaz5670

BATTERIES

How lithium dendrites form in liquid batteries

Studies of interfacial reactions and mass transport may allow safe use of lithium metal anodes

By **Jie Xiao**

Conventional rechargeable lithium (Li)–ion batteries generally use graphite as the anode, where Li ions are stored in the layered graphite. However, the use of Li metal as the anode is now being reconsidered. These next-generation battery technologies could potentially double the cell energy of conventional Li-ion batteries (*1*). Rechargeable Li metal batteries were commercialized more than four decades ago but were in use only briefly because of safety concerns (*2*). With the advancements of electrolyte (*3*, *4*), electrode architecture (*5*), and characterization techniques (*6*) in recent years, a better fundamental understanding of the interfacial reactions during charging and discharging that dictate cell performance has developed and inspired a reevaluation of the use of Li metal anodes in rechargeable batteries.

The main challenge of Li metal cells is that during charging, the Li metal electrochemically plates in an irregular manner, forming spiky microstructures, like other metals electroplated from solution, in the absence of "levelers" or "brighteners." Without these added organic compounds, whose presence results in a smoother, brighter metal surface, the metals are always dendritic or powdery (*7*). Metal dendrite formation is rooted in the mass transport of the metal cations, which are surrounded by solvent molecules and must move from bulk electrolyte to the outer limit of the electrical double layer near the electrode, followed by electro-adsorption. The solvated cations then shed the solvent molecules and are reduced into adsorbed atoms (adatoms) on the electrode surface. These adatoms diffuse on the surface and become incorporated by the metal lattice.

Mass transfer of metal cations in the electrolyte phase largely determines the final morphologies of electroplated metals, even if this process is often ignored in discussions of Li dendrite formation. Three forms of mass transport affect the cations in solution, namely diffusion, convection, and migration (see the figure). During electroreduction, the direction of cation diffusion aligns with their electromigration pathway. However, natural convection is unavoidable and unpredictable even in a static electrochemical cell (one with no net flow) and interferes with this process. Thus, some cations will move faster or slower than others, which creates different concentration gradients near the electrode (see the figure, top).

Assuming that the electrochemical deposition rate of cations is not very fast and stays the same throughout the entire electrode and that there is no interface layer formed between electrode and liquid electrolyte, a very slow movement of cations to the electrode surface can make the concentration gradient even steeper because cations are not fully replenished immediately after electroplating (see the figure, middle). Metal dendrites propagate into bulk electrolyte where more cations are available. The metal protrusion also experiences higher current densities, which self-accelerates the dendrite to grow.

If the cations move very fast in the electrolyte, the concentration gradient near the electrode is shallower (see the figure, bottom), and the metal has no preferred growth direction. Relatively large particles without sharp protrusion are usually formed when mass transport is not a concern. Once an electroplated metal particle becomes sufficiently large, it acts as a new current collector that can grow its own dendrites on the surface (*8*). The observation of randomly formed fibrous Li extending out from a Li particle reflects the different convection conditions within that area during Li plating. When a high current density is applied, electrochemical reduction or the consumption rate of Li ions is largely accelerated. Thus, the diffusion rate of Li ions in the electrolyte becomes relatively slower compared to their rate of being consumed. Strong concentration gradients are easily established throughout the entire electrode, so the metal deposition is usually highly inhomogeneous at high current densities.

Because Li metal reacts with the organic solvent of the electrolyte, the decomposition products form solid-electrolyte interphase (SEI) layers, and the higher surface areas of dendritic Li are even more reactive.

Pacific Northwest National Laboratory, Richland, WA 99352, USA. Email: jie.xiao@pnnl.gov

Insulating SEI layers worsen the uneven distribution of the electrical field within the anode and induce more dendrites to grow. The reaction between electrolyte and dendritic Li consumes electrolyte irreversibly, and the accumulation of insulating SEI layers produces "dead" Li that completely loses electronic contact with the anode (*9*). Direct stripping of Li "roots" during discharge can also disconnect and detach Li particles from the anode surface.

In cells, electrolyte and Li metal sources are both extremely limited (one magnitude less than in lab coin-cell tests). Both electrolyte and Li are quickly used up because of continuous side reactions between the two, and this degradation process sometimes occurs before short-circuiting caused by dendrite formation (*9*). Unlike Li-ion batteries, which usually display a gradual capacity that fades upon cycling, a sudden capacity drop (to almost zero) without any warning is often seen in Li metal batteries, an indicator of Li or electrolyte drainage or SEI-layer impedance buildup, or both. If a Li dendrite continues to grow without interruption, it will eventually penetrate the separator and cause an internal short. To induce the growth of a Li dendrite strong enough to short the cell, a high current density itself is insufficient. The total time of applying such a high current density also matters to short the cell.

An effective methodology needs to be developed to decouple the shorting phenomenon with the short cycle life of Li metal. One approach to promote the homogeneous electroplating of Li metal is to increase electrolyte concentration, which helps smoothen the average concentration gradients near the electrode surface (*10*). Not all concentrated electrolytes will work because the quality of SEI layers, such as the ionic conductivity and film impedance of itself derived from the electrolyte, also matters. The high viscosity of concentrated electrolyte will also limit its practical applications. A porous SEI layer should be avoided because it will retard ion diffusion toward the metal surface, accelerating establishment of a concentration gradient. The recently reported localized concentrated electrolyte contains solvent that reacts with Li to form a protective film (*11*). This reaction helped reduce unnecessary contact (and side reactions) between Li and solvent molecules.

In flow batteries (*12*), accelerated ion transport in the bulk electrolyte leads to the deposition of Li metal with much larger particle sizes (lower surface area). Pulse plating (*13*) provides time to relax the diffusion layer and slow down the evolution of a strong concentration gradient. Incorporation of a three-dimensional conductive matrix to host a Li metal anode decreases the real current density used for reducing Li ions (*14*). The carbon matrix provides a continuous electron percolation pathway that greatly slows down the formation of dead Li. However, the weight and porosity of the carbon host itself must be carefully considered to minimize the additional parasitic weight from the carbon and the excessive amount of electrolyte stored in the host.

The application of external pressure on pouch cells can extend the reversible cycling of Li metal cells. This effect is intrinsically related with the reduced amount of disconnected dead Li particles (*15*). The external pressure also improves the wetting of fresh Li surfaces by the very limited amount of electrolyte in the realistic cells, in that there is no other driving force for the lean electrolyte to flow into the new surfaces that form within porous Li structures.

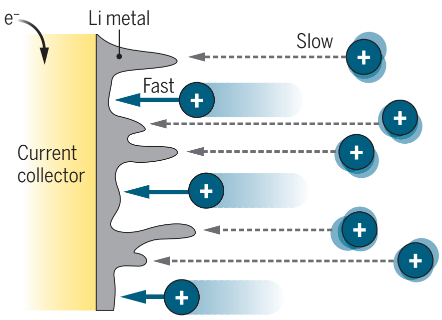
Given these fundamental considerations, it is unlikely that any approach aimed at healing dendritic Li that does not consider SEI layers will work. A very thin ionic conductive layer coated on a relatively thick Li anode will eventually fluctuate and become woven into the porous Li structures after repeated cycling. An ideal electrolyte should have little or no side reactions with Li, and thus no SEI layer is formed or needed by consuming electrolyte. Even so, mass transfer–controlled dendrite formation still exists. The electrolyte should also inhibit the formation of a strong concentration gradient near the electrode surface and promote fast diffusion of Li ions in the liquid phase. Additives that can homogenize the concentration gradient across electrode surfaces will be helpful, but they need to be reversibly usable. Approaches that reactivate dead Li will also be of interest in promoting Li anode cycling. ■

Electrodeposition of lithium (Li) ions

Mass transport of Li ions by diffusion, convection, and electromigration govern how Li metal is deposited during battery charging. The description below assumes there is no interface layer and the electroplating rate of Li is the same across the entire electrode.

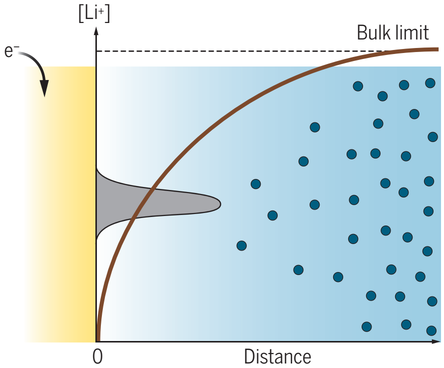
Plating unevenly

After electrodeposition, the surface of the Li metal is uneven because, locally, the mass transfer rate of Li ions (blue) varies.

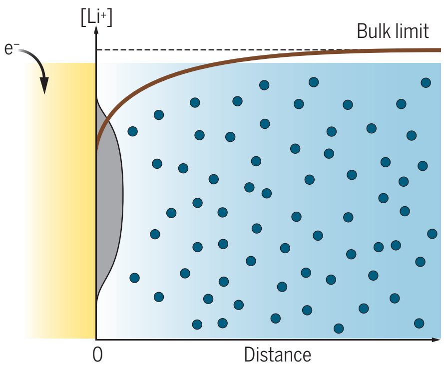


A matter of ion concentration

Schematic plots are shown for gradients in Li ion concentration, [Li⁺], and corresponding Li metal morphologies near the electrode surface.



Strong concentration gradients
Slow movement of Li ions toward the electrode causes dendrites of Li metal to grow toward the bulk electrolyte where more ions ("nutritions") are available.



Mild concentration gradients
Fast Li ion transfer provides sufficient cations in all directions and creates a smooth surface.

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

REFERENCES AND NOTES

1. J. Liu *et al.*, *Nat. Energy* **4**, 180 (2019).
2. M. S. Whittingham, *Science* **192**, 1126 (1976).
3. L. Suo *et al.*, *Science* **350**, 938 (2015).
4. Y. Yamada *et al.*, *J. Am. Chem. Soc.* **136**, 5039 (2014).
5. C. J. Bae, C. K. Erdonmez, J. W. Halloran, Y.-M. Chiang, *Adv. Mater.* **25**, 1254 (2013).
6. Y. Li *et al.*, *Science* **358**, 506 (2017).
7. L. Oniciu, L. Mureşan, *J. Appl. Electrochem.* **21**, 565 (1991).
8. X. Ai, *Energy Storage Sci. Technol.* **7**, 37 (2018).
9. C. Niu *et al.*, *Nat. Energy* **4**, 551 (2019).
10. P. Bai, J. Li, F. R. Brushett, M. Z. Bazant, *Energy Environ. Sci.* **9**, 3221 (2016).
11. S. Chen *et al.*, *Joule* **2**, 1548 (2018).
12. H. Pan *et al.*, *Adv. Energy Mater.* **5**, 1500113 (2015).
13. M. Z. Mayers *et al.*, *J. Phys. Chem.* **116**, 26214 (2012).
14. C. Niu *et al.*, *Nat. Nanotechnol.* **14**, 594 (2019).
15. A. J. Louli *et al.*, *J. Electrochem. Soc.* **166**, A1291 (2019).

ACKNOWLEDGMENTS

This research was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Vehicle Technologies of the U.S. Department of Energy through the Advanced Battery Materials Research (BMR) Program (Battery500 Consortium). Pacific Northwest National Laboratory is operated by Battelle under contract no. DE-AC05-76RL01830 for the U.S. Department of Energy.

10.1126/science.aay8672

POLICY FORUM

SOCIAL SCIENCE

Predict science to improve science

Systematic collection of predictions of research findings can provide many benefits

By **Stefano DellaVigna[1], Devin Pope[2], Eva Vivalt[3]**

Many fields of research, such as economics, psychology, political science, and medicine, have seen growing interest in new research designs to improve the rigor and credibility of research (e.g., natural experiments, lab experiments, and randomized controlled trials). Interest has similarly grown in efforts to increase transparency, such as preregistration of hypotheses and methods, that seek to allay concerns that improved research designs do not address per se, such as publication bias and p-hacking. Yet, although these efforts improve the informativeness and interpretation of research results, relatively little attention has been paid to another practice that could help to achieve this goal: relating research findings to the views of the scientific community, policy-makers, and the general public. We suggest below three broad ways in which systematic collection of predictions of research results will prove useful: by improving the interpretation of research results, mitigating bias against null results, and improving predictive accuracy and experimental design.

To date, only a relatively small number of studies have collected predictions of research, including recent work predicting original results (*1*, *2*) and the replication of academic studies (*3*–*5*). The limited attention paid to predictions of research results stands in contrast to a vast literature in the social sciences exploring people's ability to make predictions in general (*6*–*8*), as well as specifically about macroeconomic variables, geopolitical events (*9*), and sporting and political outcomes (*10*), among other variables.

We stress three main motivations for a more systematic collection of predictions of research results. The first ties to the nature of scientific progress. A new result builds on the consensus, or lack thereof, in an area and is often evaluated for how surprising, or not, it is. In turn, the novel result will lead to an updating of views. Yet we do not have a systematic procedure to capture the scientific views prior to a study, nor the updating that takes place afterward. What did people predict the study would find? How would knowing this result affect the prediction of findings of future, related studies?

Of course, informally, people routinely evaluate the novelty of scientific results with respect to what is known. However, they typically do so ex post, once the results of the new study are known. Unfortunately, once the results are known, hindsight bias ("I knew that already!") makes it difficult for researchers to truthfully reveal what they thought the results would be. This stresses the importance of collecting systematic predictions of results ex ante.

"...hindsight bias ("I knew that already!") makes it difficult for researchers to truthfully reveal what they thought the results would be."

For example, consider the debate surrounding the effectiveness of different behavioral factors and nudges to motivate a behavior. Would a gift be more, or less, effective than a modest monetary incentive? To answer this and related questions, an experiment tested how 18 different behavioral and incentive treatments (e.g., gifts, social norms, financial incentives) would motivate participants in a simple task. Notably, the researchers obtained ex ante predictions of the results for each treatment from academic experts and other forecasters such as college students (*11*, *12*).

On average, the experts were highly accurate. Furthermore, the rich data set allowed the authors to explore various features of the predictions, such as the strength of the "wisdom of the crowd" phenomenon and the relative accuracy of forecasters with different types of expertise. For example, in this context, highly cited faculty performed no better than other faculty, and Ph.D. students did best.

Another study provides an example of how predictions can be used to examine and improve belief updating based on research results, in this case in a policy setting. A group of policy-makers made predictions on the effects of conditional cash transfer programs and school meals programs (*13*). Views of policy-makers are of particular interest both because they propose and oversee interventions and because they are the people who would presumably learn from, and use, the results. Policy-makers were found to be more optimistic, but less certain, than researchers and practitioners. Further, policy-makers, practitioners, and researchers all were found to update more on "good" surprising news than on "bad" news and to not respond very differently to results with large confidence intervals as compared to results with small confidence intervals, though there is some evidence that updating can be improved by presenting results differently (*14*).

A second benefit of collecting predictions is that they can not only reveal when results are an important departure from expectations of the research community and improve the interpretation of research results, but they can also potentially help to mitigate publication bias. It is not uncommon for research findings to be met by claims that they are not surprising. This may be particularly true when researchers find null results, which are rarely published even when authors have used rigorous methods to answer important questions (*15*). However, if priors are collected before carrying out a study, the results can be compared to the average expert prediction, rather than to the null hypothesis of no effect. This would allow researchers to confirm that some results were unexpected, potentially making them more interesting and informative, because they indicate rejection of a prior held by the research community; this could contribute to alleviating publication bias against null results.

A third benefit of collecting predictions systematically is that it makes it possible

[1]Department of Economics, University of California, Berkeley, Berkeley, CA, USA. [2]The University of Chicago Booth School of Business, Chicago, IL, USA. [3]Research School of Economics, Australian National University, Canberra, Australia. Email: eva.vivalt@anu.edu.au

Prediction platform

Process to collect forecasts for comparison with study results

1. Researcher designs study and collects baseline data if applicable.
2. Researcher designs forecasting survey and sends it to the platform.
3. The platform distributes the forecasting survey.
4. Researcher gathers results data for the study.
5. Forecasting survey results are released to researcher at prespecified date.
6. Study results are released back to forecasters at end of study.
7. Optional follow-up survey is conducted with forecasters to measure belief updating.

Sample outcomes for forecasting survey

What is the increase (in standard deviation units) of savings in the treatment group, compared to in the control group?

Did the program cause employment rates to increase?

What are your predictions about the average number of points scored in each of the 15 remaining conditions?

Sample summary statistics to elicit:

The mean effect

A range of values such that the respondent is 90% sure the mean effect falls within that range

Whether the study result will be positive and significant, insignificant, or negative and significant

to improve the accuracy of predictions. In turn, this may help with experimental design. For example, envision a behavioral research team consulted to help a city recruit a more diverse police department. The team has a dozen ideas for reaching out to minority applicants, but the sample size allows for only three treatments to be tested with adequate statistical power. Fortunately, the team has recorded forecasts for several years, keeping track of predictive accuracy, and they have learned that they can combine team members' predictions, giving more weight to "superforecasters" (*9*). Informed by its longitudinal data on forecasts, the team can elicit predictions for each potential project and weed out those interventions judged to have a low chance of success or focus on those interventions with a higher value of information. In addition, the research results of those projects that did go forward would be more impactful if accompanied by predictions that allow better interpretation of results in light of the conventional wisdom.

These three broad uses of predictions highlight two important implications. First, it will be important to collect forecast data systematically to draw general lessons. For example, when do senior researchers make more accurate forecasts than junior researchers, given that the seniors' expertise did not help in the study and forecasts of task performance incentives (*12*)? Under what conditions do policy-makers update in a Bayesian manner from past evidence (*13*)? We will need predictions for a range of settings, including longitudinal predictions by the same forecasters over time, to identify possible superforecasters and to examine whether providing feedback on past forecasts helps improve prediction accuracy.

Second, like preanalysis plans, it is critical to set up the collection of predictions before the results are known, to avoid the impact of hindsight bias. With these features in mind, a centralized platform that collects forecasts of future research results can play an important role. Toward this end, in coordination with the Berkeley Initiative for Transparency in the Social Sciences (BITSS), we have developed an online platform for collecting forecasts of social science research results (www.socialscienceprediction.org). The platform will make it possible to track multiple forecasts for an individual across a variety of interventions, and thus to study determinants of forecast accuracy, such as characteristics of forecasters or interventions, and to identify superforecasters (see the box).

A centralized platform has another advantage. As collecting forecasts grows in popularity, a small number of researchers may receive a disproportionate number of requests. A centralized platform can ensure that this does not happen, analogous to how an editor keeps track of referee requests within a journal, except that a centralized platform could be even better, as editors cannot track referee requests across journals. As a further benefit, the platform provides third-party certification about how forecasts were collected and shared with researchers requesting them (analogous to platforms used for preregistration).

This platform would aim to incorporate lessons learned from other work on forecasts, such as work on replication of experiments in psychology and economics, prediction of geopolitical events in the Good Judgment Project, and forecasts of macroeconomic indicators in the Survey of Professional Forecasters. The Systematizing Confidence in Open Research and Evidence (SCORE) program is also aiming to develop tools specifically to predict the replicability or reproducibility of social-behavioral science results.

There are many open questions about the details of the platform. For example, should forecasters be paid for participating (just like some journals choose to pay referees)? Should there be incentives for accuracy? We expect that continued work and experimentation will provide more clarity regarding such design questions.

Although here we focus on the benefits of ex ante predictions for improving the interpretation of research results, these predictions have many other potential uses in research and policy. Some researchers may use predictions to explore when forecasts can be trusted or how the accuracy of forecasts can be improved. Others will focus on Bayesian interpretations or learning about belief updating. The forecasts may also have a practical value to policy-makers needing to make a decision in the absence of credible evidence from an academic study. Such a variety of potential uses speaks to the value of making this tool available. ■

REFERENCES AND NOTES

1. A. Delavande, X. Giné, D. McKenzie, *J. Dev. Econ.* **94**, 151 (2011).
2. A. Cohn, M. A. Maréchal, D. Tannenbaum, C. L. Zünd, *Science* **365**, 70 (2019).
3. A. Dreber *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15343 (2015).
4. C. F. Camerer *et al.*, *Science* **351**, 1433 (2016).
5. D. Benjamin, D. R. Mandel, J. Kimmelman, *PLOS Biol.* **15**, e2002212 (2017).
6. F. Galton, *Nature* **75**, 450 (1907).
7. D. Kahneman, A. Tversky, *Psychol. Rev.* **80**, 237 (1973).
8. R. M. Dawes, D. Faust, P. E. Meehl, *Science* **243**, 1668 (1989).
9. P. Tetlock, D. Gardner, *Superforecasting: The Art and Science of Prediction* (Random House, 2016).
10. J. Wolfers, E. Zitzewitz, *J. Econ. Perspect.* **18**, 107 (2004).
11. S. DellaVigna, D. Pope, *Rev. Econ. Stud.* **85**, 1029 (2018).
12. S. DellaVigna, D. Pope, *J. Polit. Econ.* **126**, 2410 (2018).
13. E. Vivalt, A. Coville, "How Do Policymakers Update?" Working paper (2016).
14. E. Vivalt, A. Coville, "The Implications of Variance Neglect for the Formation and Estimation of Subjective Expectations." Working paper (2019).
15. A. Franco, N. Malhotra, G. Simonovits, *Science* **345**, 1502 (2014).

ACKNOWLEDGMENTS

We thank A. Coville, D. Karlan, E. Miguel, N. Otis, and participants at the Forecasting Results in the Social Sciences Workshop at the University of California, Berkeley, for helpful discussions. We gratefully acknowledge the support of the Alfred P. Sloan Foundation in building the Social Science Prediction Platform, (G-2019-12325), and an anonymous foundation. E.V. is also supported by the John Mitchell Economics of Poverty Lab at the Australian National University.

10.1126/science.aaz1704

The remains at Colorado's Corral Bluffs, including these four skulls, reveal much about early mammals.

PALEONTOLOGY

Stepping out of the dinosaurian shadow

A new film offers a glimpse into the Cenozoic ascension of mammals

By **Riley Black**

We live in the Age of Mammals, yet warm-blooded beasts are still overshadowed by dinosaurs. Even when considering the last great shake-up to life's story, when an enormous asteroid triggered a mass extinction that decimated dinosaurs and gave mammals a shot at terrestrial expansion, we are often more focused on the terrible lizards lost than the furry creatures who set the stage for the Cenozoic. But a new NOVA documentary—"Rise of the Mammals"—seeks to change that and, in the process, offers viewers a window into paleontology beyond bone hunting.

Narrated in soothing tones by actor Keith David, the 1-hour program promises to tell how life surged back after the Cretaceous-Paleogene mass extinction. The catastrophe, we learn, did not just affect dinosaurs. Flying pterosaurs and seagoing mosasaurs disappeared, as did coil-shelled ammonites and huge clams called rudists, along with mass extinctions of birds, lizards, and mammals. But the mammal-versus-dinosaur competition is the film's primary focus, with images of tiny, shrewlike insectivores living beneath the feet of dinosaurs providing the background for what follows.

The "Mesozoic mammals as underdogs" trope should be extinct by now. In the past several decades, paleontologists have recognized that mammals thrived during this era, evolving into an impressive array of forms. All were small, fair enough, but so are most of today's mammal species. To navigate a clear relationship of cause and effect, then, "Rise of the Mammals" emphasizes size. When did mammals start to get big?

An arid field site called Corral Bluffs near Colorado Springs, Colorado, is offered as the key place to answer this question. How this dot on the map was uncovered is told in a circuitous fashion. We are first introduced to Denver Museum of Nature & Science paleontologist Tyler Lyson and his quest to find fossil beds from the earliest days of the Paleocene, the epoch directly following the Cretaceous.

"Rise of the Mammals" leans heavily on the romance of fieldwork and *Indiana Jones* imagery here, even cribbing a famous sunset shot from *Raiders of the Lost Ark*. But the key to the story is not a new discovery by Lyson. It is a previous find that was already in the Denver collections.

Paleontology lore often focuses on first authors, field leaders, and museum curators to the exclusion of other workers who make the science possible. To the documentary's credit, the contribution of Denver museum volunteer Sharon Milito—who made the first critical find at Corral Bluffs, picking up a hard concretion that preserved a Paleocene mammal palate inside—is recognized and underscored by an interview with Milito herself. Lyson happened across this fossil in the Denver collections and, upon visiting the site with colleague Ian Miller, started finding dozens more well-preserved Paleocene fossils.

Rise of the Mammals
HHMI Tangled Bank Studios
Premiering 30 October 2019 on PBS at 9 pm ET

The film quickly shifts into detective mode. What species were found at the site? How old were they? What was their environment like? Instead of following the standard—and often false—story of how a single discovery changes everything, the program follows various threads to assemble a picture of life in the first million years of the Paleocene.

Although the legacy of storytelling from the "Bone Wars" era of epic fossil hunts is certainly there, the latter half of the film broadens in scope. "Rise of the Mammals" ends up being a short course in modern paleontology, following the story as it goes back and forth between museum and fossil outcrops. The changing face of paleontology is visible, too; the cast of scientists shown and interviewed is much more gender-balanced and diverse than many programs of the past few decades.

The Corral Bluffs fossils are phenomenal, and what they have to tell us about the Paleocene is just starting to drip out into the published record (*1*), but the ancient ecosystem is only one small part of a global story. Where the film shines—and offers something rare—are the moments when the process of science is allowed to unfold, revealing how experts assemble views of lost worlds. And if nothing else, it is helpful to pry the spotlight out of dinosaurian claws now and then. ■

The reviewer is the author of *Skeleton Keys: The Secret Life of Bone* (Riverhead Books, 2019) and *My Beloved Brontosaurus: On the Road with Old Bones, New Science, and Our Favorite Dinosaurs* (Scientific American/Farrar, Straus and Giroux, 2014). Email: evogeek@gmail.com

REFERENCES AND NOTES

1. T. R. Lyson *et al.*, Science 10.1126/science.aay2268 (2019).

10.1126/science.aaz6313

LETTERS

Army troops helped protect the greater one-horned rhino population in Chitwan National Park, Nepal.

PHOTO: BIOSPHOTO/ALAMY STOCK PHOTO

Edited by **Jennifer Sills**

Editor's note

When the Letter "Trophy hunting bans imperil biodiversity" (A. Dickman *et al.*, 30 August, p. 874) was published, *Science*'s policy of asking all manuscript authors to declare conflicts of interest did not apply to Letters. This policy is now under revision to ensure that authors of Letters also make readers aware of financial and advisory competing interests. *Science* has therefore requested that the authors of Dickman *et al.* declare their competing interests. They have done so in an addendum to their Letter.

Jeremy Berg
Editor-in-Chief

10.1126/science.aaz9111

Trophy hunting: Role of consequentialism

In their Letter "Trophy hunting bans imperil biodiversity" (30 August, p. 874), A. Dickman *et al.* adopt a radical consequentialist approach. Not only do they discard any deontological concern relevant to trophy hunting but also, remarkably, they oppose policies that would consider deontological objections against trophy hunting. Said otherwise, according to Dickman *et al.*, evidence-based policy-making must trump moral-based policy-making.

Consequentialist approaches are not uncommon in conservation (*1*), but Dickman *et al.*'s Letter is important because it opens the question of whether ethical objections limit nature conservation and whether it is time to move beyond such objections, as the authors do. Conservation increasingly operates within an ethical frame whereby protecting wild plant and animal species must first and foremost benefit human communities and becomes unacceptable if it imposes a burden on people. This emphasis is leveraged by social science, which has a growing importance in conservation and is becoming more concerned with social justice than with an objective understanding of social systems (*2*). As a result, conservation practices such as green militarization or human population displacement are often arbitrarily excluded by scholars from the conservation toolkit and mostly mentioned from a critical or adversarial ideological standpoint (*3*).

Yet, these practices have the potential to deliver impressive results. The greater one-horned rhino (*Rhinoceros unicornis*) population in Chitwan National Park, Nepal rebounded after the deployment of army troops to fight against poaching and the expansion of the park (*4*). African Parks—a nongovernmental organization managing protected areas with a total control approach—set up anti-poaching forces with SWAT-like training and succeeded in increasing Chad's Zakouma National Park elephant populations (*5*). In the Central African Republic, African Parks purchased from Bulgaria more than one hundred war-grade weapons with 90,000 rounds under an exemption from the UN embargo and shipped them to the Chinko Project, a 17,600-km^2 wildlife refuge it manages with full law enforcement competence (*6*). Displacement is another ostracized conservation tool (*7*). However, by displacing thousands of people in northeastern China, the government has reduced human population density by more than half and consequently Amur tigers (*Panthera tigris altaica*) and leopards (*Panthera pardus orientalis*) are recovering in the area (*8*).

Opposing green militarization or population displacement negatively affects conservation, and viable alternatives are often lacking. Although some people find these approaches unethical, conservation policy that is not based on science threatens habitat and biodiversity. If ethical concerns remain selective and subjective, conservation is unlikely to succeed. The same pragmatic, results-oriented rationale that Dickman *et al.* advocate for trophy hunting may need to be expanded to other controversial conservation approaches.

Guillaume Chapron[1,2]* and José Vicente López-Bao[3]
[1]Department of Ecology, Swedish University of Agricultural Sciences, 730 91 Riddarhyttan, Sweden. [2]Wildlife Conservation Research Unit, Department of Zoology, University of Oxford, Tubney, OX13 5QL, UK. [3]Research Unit of Biodiversity (UO/CSIC/PA), Oviedo University, 33600 Mieres, Spain.
*Corresponding author. Email: guillaume.chapron@slu.se

REFERENCES AND NOTES

1. A. Hiller, R. Ilea, L. Kahn, *Consequentialism and Environmental Ethics: Routledge Studies in Ethics and Moral Theory* (Taylor & Francis, 2013).
2. J. Toby, "Left-wing politics and the decline of sociology," *The Wall Street Journal* (2019).
3. R. Duffy *et al.*, *Biol. Conserv*. **232**, 66 (2019).
4. A. Dudley, "Nepal's rhino numbers rise, thanks to national and local commitment," *Mongabay* (2017).
5. R. Nuwer, "The rare African park where elephants are thriving," *National Geographic* (2017).
6. UN Security Council, S/2018/729 Overview of Sanctions Committee Documents (2018).
7. C. Geisler, *Int. Soc. Sci. J.* **55**, 69 (2003).
8. G. Jiang *et al.*, *Biol. Conserv.* **211**, 142 (2017).

COMPETING INTERESTS

G.C. is a member of the IUCN Large Carnivore Initiative for Europe, IUCN Canid Specialist Group, and IUCN Cat Specialist Group. J.V.L.-B. is a member of the IUCN Canid Specialist Group.

10.1126/science.aaz4951

Trophy hunting: Values inform policy

In their Letter "Trophy hunting bans imperil biodiversity" (30 August, p. 874), A. Dickman *et al.* mischaracterize context, offer weak evidence, and overlook the role of values. They caution against trophy hunting bans, yet the policies they cite do not ban trophy hunting. Two of the policies discontinue only import of lion trophies (*1*); the others ban the import of trophies from a delimited set of endangered species (*1–3*). These are not blanket bans on trophy hunting but species-specific import restrictions. Although Dickman *et al.* contend such bans would "imperil biodiversity," their evidence is selective [e.g., (*4*, *5*)] and does not directly support the contention that import bans yield negative conservation outcomes.

In raising concerns about sustainable community development, Dickman *et al.* set up a false dichotomy: Either restrict the import of wildlife trophies to Western countries or promote self-sustaining African communities. Western nations can support sustainable development of African nations while regulating the import of wildlife trophies by their own citizens. Although import bans in Western nations potentially affect African communities, these impacts should not be confused with the impacts of discontinuing trophy hunting. Especially where trophy hunting generates few benefits for local people (*6*, *7*), negative socioeconomic effects of import bans will likely be limited.

Dickman *et al.* further assert that "calls for hunting bans usually cite conservation concerns," but such calls are often motivated by moral concerns (*1*, *3*, *8*). Indeed, the authors allude to this in suggesting policy should be based on science, not feelings of "repugnance." This position establishes another false dichotomy. Adjudicating policy requires both understanding the likely results of a policy (science) and evaluating whether those results are desirable (values) (*9*). Such evaluative judgments are expressed by emotions (*10*). Policies supporting sustainable community development may seek to remediate the harms of Africa's colonial history. Recognizing these harms as injustices—a moral judgment—engenders emotions such as anger. Policies may also aim to combat perceived injustice against nonhuman animals, which may similarly elicit outrage. In short, emotion attends moral judgment, which informs policy.

Conservation is rife with risk. Humans and wildlife face physical and biological risks; hence both are subjects of concern. But conservation strategies may carry moral risks as well, even when enacted out of concern. Science can quantify risks, but it cannot tell us whether they are acceptable or by whose values they should be judged. Governments are right to institute policies that manage the landscape of risk by weighing scientific evidence and accounting for the values of their citizens.

Chelsea Batavia[1]*, Jeremy T. Bruskotter[2], Chris T. Darimont[3,4], Michael Paul Nelson[1], Arian D. Wallach[5], and 56 signatories
[1]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, 97331, USA. [2]School of Environment & Natural Resources, Ohio State University, Columbus, OH 43210, USA. [3]Department of Geography, University of Victoria, Victoria, BC V8W 2Y2, Canada. [4]Raincoast Conservation Foundation, Sidney, BC V8L 3Y3 Canada. [5]Centre for Compassionate Conservation, School of Life Sciences, University of Technology Sydney, Ultimo, NSW, 2007, Australia.
*Corresponding author.
Email: chelsea.batavia@oregonstate.edu

REFERENCES AND NOTES

1. E. Ares, "Trophy hunting," Briefing Paper 7908 (House of Commons Library, London, UK, 2019).
2. U.S. Congress, H.R. 2245—CECIL Act (2019).
3. H. Horton, "Britain will have toughest trophy hunting rules in the world as Government announces ban of 'morally indefensible' act," *The Telegraph* (2019).
4. D. W. Macdonald *et al*., *Mammal Rev.* **47**, 247 (2017).
5. W. J. Ripple *et al*., *Trends Ecol. Evol.* **31**, 495 (2016).
6. T. Pasmans, P. Hebinck, *Land Use Pol.* **64**, 440 (2017).
7. A. Yasuda, *Soc. Nat. Resour.* **24**, 860 (2011).
8. Born Free Foundation, "Ban trophy hunting: Help end this barbaric sport" (2019).
9. K. D. Moore, M. P. Nelson, Eds., *Moral Ground: Ethical Action for a Planet in Peril* (Trinity University Press, San Antonio, TX, 2010).
10. A. Damasio, *Looking for Spinoza* (Houghton Mifflin Harcourt Publishing, New York, 2003).

COMPETING INTERESTS

J.T.B. has received funding from the Ohio Agricultural Research and Development Center, the Ohio Department of Natural Resources, the National Oceanic and Atmospheric Administration, and the Association for Fish and Wildlife Agencies. He serves in an advisory capacity for Project Coyote and the Ohio Department of Natural Resources. M.P.N. has received funding from the National Science Foundation and serves in an advisory capacity for Project Coyote.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/433.1/suppl/DC1
List of signatories

10.1126/science.aaz4023

Response

"…Batavia *et al.* urge us not to confuse the effect of restricting imports on local people with the impacts of discontinuing trophy hunting [but] the former is meant to bring about the latter…" —Dickman *et al.*

Full text: science.sciencemag.org/content/366/6464/433.1/tab-e-letters

Trophy hunting: Broaden the debate

In their Letter "Trophy hunting bans imperil biodiversity" (30 August, p. 874), A. Dickman *et al.* argue against trophy hunting bans, but the bans they mention are neither blanket nor hunting bans. France only suspended lion trophy imports, whereas Australia and The Netherlands banned import permits for trophies of several species (*1*), but other trophies continue to be collected worldwide and domestically. Indeed, a ban on the import of a trophy into a nation does not constitute a ban on hunting by its nationals; Dickman *et al.* confound the two, which is disingenuous and raises the question of whether hunting is a sport or a form of commodity acquisition (*2*).

International movement of trophies is regulated under the Convention on International Trade in Endangered Species (CITES), governed by member states. Unless there is evidence of trade threatening the survival of a species, sovereign states can allow hunting and export trophies, but potential importing states also have sovereignty over their response to concerns of their constituencies and have the right to implement what CITES calls "stricter domestic measures" (*3*). Debates around this have been politicized, which typically happens when scientific data are too inconclusive to guide policy formulation (*3*).

Dickman *et al.* misrepresent the responsibility of importing states over hunting policy; ironically, they may stimulate blanket bans by arguing against opt-outs for some countries for certain species. Moreover, they fail to mention that where hunting zones are protected areas recognized by civil law, they would remain so. In addition, habitat in hunting zones is often not effectively protected, and the collapse of trophy hunting observed in certain areas is not due to trade bans but to a failing balance of costs and benefits (*4*, *5*). Trophy hunting is neither the main threat to nor the main opportunity for wildlife conservation, and we encourage a broader debate.

Hans Bauer[1]*, Bertrand Chardonnet[2], Mark Jones[3], Claudio Sillero-Zubiri[1,3]
[1]Wildlife Conservation Research Unit, Zoology, University of Oxford. Tubney OX13 5QL, UK. [2]African Protected Areas and Wildlife, 92210 Saint Cloud, France. [3]Born Free Foundation, Horsham, RH12 4QP, UK.
*Corresponding author.
Email: hans.bauer@zoo.ox.ac.uk

REFERENCES AND NOTES

1. E. Ares, "Trophy hunting," Briefing Paper 7908 (House of Commons Library, London, UK, 2019).
2. C. Batavia *et al.*, *Conserv. Lett.* **12**, e12565 (2019).
3. H. Bauer *et al*., *Conserv. Lett.* **11**, e12444 (2018).
4. H. Bauer *et al.*, *PLOS One* **12**, e0173691 (2017).
5. B. Chardonnet, "Africa is changing: Should its protected areas evolve? Reconfiguring the protected areas in Africa" (IUCN, 2019).

COMPETING INTERESTS

H.B. has received funding from or is a member of University of Oxford, Born Free Foundation, IUCN Cat Specialist group, IUCN Save Our Species, Wildlife Conservation Network, National Geographic, U.S. Fish and Wildlife Service, Leo Foundation, GIZ (German Technical Development

Cooperation), and Leipzig Zoo. B.C. has received funding from or is a member of African Union/IBAR, IUCN, GIZ (German Technical Development Cooperation), KfW (German Financial Cooperation), Afrique Nature International, Ecole Inter-Etats des Sciences et Médecine Veterinaire de Dakar (EISMV)/University of Minnesota, African Wildlife Foundation, OIE–World Organization for Animal Health, and AFD (French Agency for Development). C.S.-Z. is Chair of the IUCN Species Survival Commission Canid Specialist Group and Chief Scientist of Born Free Foundation. He has received funding from or is a member of University of Oxford, University of Vermont, Wildlife Conservation Network, Fondation Segre, Conservation International Critical Ecosystem Partnership Fund, African Wildlife Foundation, IUCN Save Our Species, National Geographic, U.S. Fish and Wildlife Service, IUCN Cat Specialist group, IUCN Wildlife Health Specialist group, and IUCN Human Wildlife Conflict Task Force.

10.1126/science.aaz4036

Response

"...It is...hypocritical for a rich country to...reduce the viability of trophy hunting in poor countries while taking no action against domestic sport hunting..."
—Dickman *et al.*

Full text: science.sciencemag.org/content/366/6464/433.2/tab-e-letters

Trophy hunting: Bans create opening for change

In their Letter "Trophy hunting bans imperil biodiversity" (30 August, p. 874), A. Dickman *et al.* warn that banning trophy hunting, a practice many of them deem "repugnant," could threaten African biodiversity and livelihoods. What they actually describe is how loss of funding may impart these effects, without specifying any unique benefits of trophy hunting. It is defeatist to defend business-as-usual instead of promoting alternative conservation activities that could sustain formerly trophy-hunted species and areas.

Trophy hunting relies on deep geopolitical inequalities, particularly in Africa, where it often fails to deliver demonstrable conservation outcomes (*1*) and can intersect with crime (*2*). It yields low returns at household levels (*3*), with only a fraction of generated income reaching local communities (*4*). It also siphons off wildlife from adjacent protected areas (*5*), reduces population connectivity and resilience, and can have genetic consequences such as reductions in body, horn, and/or tusk size (*6*). Its effects on wildlife demography and behavior can be profound (*7*).

Trophy import bans present an opportunity to rethink how we can conserve wildlife in nonextractive ways that are consistent with shifting public opinion. The system is primed for change. The recently polled U.S. public shares attitudes with other countries enacting trophy import bans and especially strongly disapproves of trophy hunting of African elephants and lions (*8*). Sustainable alternatives exist and could reduce reliance on a small and narrowing cohort of wealthy Western "donors" (*9*).

For example, land use reforms, co-management, and greater participatory stewardship can provide a more sustainable, resilient, and equitable system (*10*). Locally adjusted and bottom-up management practices (*11*), granting communities land titles, conservation-compatible agriculture, and coexistence approaches can also benefit communities and conservation more than trophy hunting. In addition, tourism reforms could invigorate domestic tourism (*12*), minimize leakage of tourism income to foreign investors, and reduce the footprint of wildlife-viewing tourism through green development investment. Diversified nature-based tourism beyond photographing and viewing wildlife could incorporate survival skills/bushcraft training and agritourism, emphasizing local knowledge, cultural exchange, and inclusion of women. Finally, environmental investments could connect would-be micro-investors more directly to wildlife-wealthy communities. Financial strategies such as decentralized markets made possible by blockchain technology could use carbon and biodiversity credits for conserving habitats. Sustainable enterprise development could generate direct financial benefits to local communities.

During transitions, nongovernmental organizations could raise funds to pay concessions or countries could agree that a private entity would temporarily assume game reserve management. As the bans are not blanket but import bans, they provide the impetus and the time to incrementally switch to practices that maximize contributions to the Sustainable Development Goals.

Katarzyna Nowak[1,2]*, Phyllis C. Lee[3,4], Jorgelina Marino[5], Mucha Mkono[6], Hannah Mumby[7,8,9], Andrew Dobson[10], Ross Harvey[11], Keith Lindsay[4], David Lusseau[12], Claudio Sillero-Zubiri[5,13], and 71 signatories

[1]The Safina Center, Setauket, NY 11733, USA. [2]Department of Zoology and Entomology, University of the Free State, Phuthaditjhaba, 9866, South Africa. [3]Faculty of Natural Sciences, University of Stirling, Stirling, FK9 4LA, UK. [4]Amboseli Trust for Elephants, Langata, Nairobi 00509, Kenya. [5]Wildlife Conservation Research Unit, Zoology, University of Oxford, Tubney OX13 5QL, UK. [6]Tourism Cluster, University of Queensland Business School, University of Queensland, QLD 4072, Australia. [7]School of Biological Sciences, University of Hong Kong, Pok Fu Lam, Hong Kong. [8]Department of Politics and Public Administration, University of Hong Kong, Pok Fu Lam, Hong Kong. [9]Centre for African Ecology, School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, 2000, South Africa. [10]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA. [11]School of Economics, University of Cape Town, Cape Town, South Africa. [12]School of Biological Sciences, University of Aberdeen, Aberdeen AB24 2TZ, UK. [13]Born Free Foundation, Horsham, RH12 4QP, UK.
*Corresponding author.
Email: knowak02@gmail.com

REFERENCES AND NOTES

1. J. Selier *et al.*, *J. Wildlife Manage.* **78**, 122 (2014).
2. T. Milliken, J. Shaw, "The South Africa–Viet Nam Rhino Horn Trade Nexus" (TRAFFIC, 2012).
3. M. Segage, Master's thesis, University of Limpopo (2015).
4. I. *Nordbø et al.*, *J. Sustain. Tour.* ***26**, 68 (2018).*
5. A. J. Loveridge *et al.*, *Biol. Conserv.* **134**, 548 (2016).
6. D. W. Coltman *et al.*, *Nature* **426**, 655 (2003).
7. J. M. Milner *et al.*, *Conserv. Biol.* **21**, 36 (2007).
8. Responsive Management, "Americans' attitudes toward hunting, fishing, sport shooting and trapping" (NSSF, 2019).
9. C. Batavia *et al.*, *Conserv. Lett.* **12**, e12565 (2018).
10. IPBES, "IPBES global assessment summary for policy-makers" (2019).
11. Z. T. Ashenafi, N. Leader-Williams, *Hum. Ecol.* **33**, 539 (2005).
12. S. B. Mariki *et al.*, *J. Environ. Pol. Plan.* **4**, 62 (2011).

Tourism reforms could make wildlife-viewing tourism greener and more beneficial to local communities.

PHOTO: MARKA/CONTRIBUTOR/GETTY IMAGES

COMPETING INTERESTS

P.C.L. is Director of Science, Amboseli Trust for Elephants. R.H. consults for the Conservation Action Trust and the EMS Foundation. D.L. is a member of IUCN Sustainable Use and Livelihoods Specialist Group and a member of IUCN Species Survival Commission Cetacean Specialist Group. C.S.-Z. is Chair of the IUCN Species Survival Commission Canid Specialist Group and Chief Scientist of Born Free Foundation. He has received funding from or is a member of University of Oxford, University of Vermont, Wildlife Conservation Network, Fondation Segre, Conservation International Critical Ecosystem Partnership Fund, African Wildlife Foundation, IUCN Save Our Species, National Geographic, U.S. Fish and Wildlife Service, IUCN Cat Specialist group, IUCN Wildlife Health Specialist group, and IUCN Human Wildlife Conflict Task Force.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/434/suppl/DC1
List of signatories

10.1126/science.aaz4135

Response

"...[T]he true risk is...losing funding streams that require the presence of trophy hunted species...and therefore incentivize conservation of their populations and habitat..." —Dickman *et al*.

Full text: science.sciencemag.org/content/366/6464/434/tab-e-letters

Trophy hunting: Insufficient evidence

In their Letter "Trophy hunting bans imperil biodiversity" (30 August, p. 874), A. Dickman *et al*. argue that banning trophy hunting would be detrimental to conservation. We agree that evidence for effectiveness is important before actions are taken. However, Dickman *et al*. do not provide evidence that bans to trophy hunting harm biodiversity (*1*).

Dickman *et al*. claim that trophy hunting indirectly benefits biodiversity because populations (and their habitats) are better protected in places or times where trophy hunting has occurred. However, no comprehensive research has tested that hypothesis. Even previous work by Letter authors Dickman and Johnson (led by Macdonald) concludes that we know too little to infer whether trophy hunting (selective hunting for recreation) contributes to the conservation of wild lions (*2*)—one of the best-studied trophy-hunted species.

Dickman *et al*. overstate their case. For example, the claim that "more land has been conserved under trophy hunting than under national parks" seems based on the statement from Lindsey *et al*. (*3*) that "[o]ver 1,394,000 km^2 is used for hunting in sub-Saharan Africa, exceeding the area encompassed by national parks by 22% in the countries where hunting is permitted" (*3*). However, this interpretation is misleading because those lands include private lands, protected areas that allowed subsistence hunting, and various other classes of protected areas, not exclusively trophy hunting concessions. In addition, the authors' prediction that a ban on trophy imports or hunts would indirectly harm biodiversity could be just the converse: Perhaps hunting concessions would be upgraded in protection by catalyzing a governmental rethinking of carnivore management systems. An evidentiary basis for informing controversial policy interventions, such as trophy hunting, demands strong inference with full disclosure of uncertainties and disentangled value judgments from observations or inferences.

Stronger evidence might be gleaned through adequate tests of the effectiveness of trophy hunting for protecting the hunted population, including broad-scale experiments using multiple replicated land parcels subject either to hunting or another putative form of biodiversity protection under similar socioeconomic systems, or tracking of populations before and after trophy hunting (accounting for other threats). Rigorous examinations would likely reveal outcomes that vary by population, geography, other threats to biodiversity, and socioeconomic and governance contexts.

Finally, the addition of a long list of signatories implies a call to authority that should play little or no role in what should ultimately be an evidence-based scientific debate. By contrast, clear evidence, transparently conveyed and clearly demarcated from the ingrained values of those involved (whether they support or reject trophy hunting), could help elucidate environmental, ethical, social, and economic dimensions of this controversial activity whose ultimate conservation effects remain poorly understood.

A. Treves[1]*, F. J. Santiago-Ávila[1], V. D. Popescu[2,3], P. C. Paquet[4,5], W. S. Lynn[6,7], C. T. Darimont[4,5], K. A. Artelle[4,5]

[1]Nelson Institute for Environmental studies, University of Wisconsin, Madison, WI 53706 USA. [2]Department of Biological Sciences and Sustainability Studies Theme, Ohio University, Athens, OH 45701, USA. [3]Center for Environmental Research (CCMESI), University of Bucharest, Bucharest, Romania. [4]Department of Geography, University of Victoria, Victoria, BC V8W 2Y2, Canada. [5]Raincoast Conservation Foundation, Sidney, BC V8L 3Y3, Canada. [6]Marsh Institute, Clark University, Worcester, MA 01610, USA. [7]Knology, New York, NY 10005, USA.
*Corresponding author. Email: atreves@wisc.edu

REFERENCES AND NOTES

1. A. Treves *et al*., *Conserv. Biol.* **33**, 472 (2018).
2. D. W. Macdonald *et al*., *Mamm. Rev.* **47**, 247 (2017).
3. P. A. Lindsey *et al*., *Biol. Conserv.* **134**, 455 (2007).

COMPETING INTERESTS

A.T. is President of the Board of Directors of Future Wildlife, a tax-exempt organization with the mission to preserve nature, especially wild animals, and an unpaid science adviser for Project Coyote.

10.1126/science.aaz4389

Response

"...[A]ction should not be taken without evidence for its effectiveness...[but] we believe the burden of proof clearly lies with those who support [the removal of trophy hunting]..." —Dickman *et al*.

Full text: science.sciencemag.org/content/366/6464/435.1/tab-e-letters

Trophy hunting: A moral imperative for bans

In their Letter "Trophy hunting bans imperil biodiversity" (30 August, p. 874), A. Dickman *et al*. argue that trophy hunting should not be discontinued. However, their premise is not viable when examined under the light of basic morality.

Whether Dickman *et al*. concur or not, wildlife has the basic right of existence, irrespective of human existence and interests. Intentional killing of animals to satisfy the whims of wealthy individuals is detestable. No potential gains, even those that are promoted by Dickman *et al*. as beneficial to wildlife, justify undermining the moral basis of the protection of Earth's natural resources. It is our responsibility to suppress the destructive tools at our disposal so that these resources remain unharmed.

Culling of endangered species is a self-evident fallacy. Our foremost emergency is to restore endangered species to their former state, irrespective of human interests. Unless required for basic existence, hunting of all forms is a practice that should be eradicated like the smallpox virus. Beyond rational arguments, the most appropriate response to the Letter by Dickman *et al*. is outrage.

Arie Horowitz
Philadelphia, PA, 19106, USA
Email: arie2006@gmail.com

10.1126/science.aaz3315

Response

"...[Discontinuing] trophy hunting...without implementing better alternatives risks worsening the situation for both wildlife and people..." —Dickman *et al*.

Full text: science.sciencemag.org/content/366/6464/435.2/tab-e-letters

AAAS NEWS & NOTES

An Oregon shellfish hatchery successfully monitors seawater to rejuvenate oyster larvae, reviving oyster population levels and restoring a Pacific Coast industry.

AAAS's "How We Respond" report captures U.S. ingenuity

Facing climate change, communities across the country devise locally tailored responses, report shows

By **Anne Q. Hoy**

Communities from Arizona to Montana and Massachusetts to California are recognizing the urgency to take steps to limit the impact of climate change. A two-year project by the American Association for the Advancement of Science spotlights examples of 18 innovative solutions communities are putting in place to blunt shifts in Earth's natural systems.

Kansas farmers, for instance, have cut back on water use to preserve groundwater resources in the nation's increasingly arid Southern Great Plains region.

An avid Texas churchgoer established an auditing company to improve energy efficiency at some 500 churches, freeing money for charitable endeavors.

Wisconsin's Dane County has been capturing methane gas from its largest landfill, converting it into renewable transportation fuel, and reducing its greenhouse gas emissions.

Rising death rates of oyster larvae at a northern Oregon shellfish hatchery led to a regional network of seawater monitoring tools able to measure acidity levels and restore oyster populations, salvaging a Netarts Bay and upper Pacific Coast industry.

Threatened by storm surge from perilous hurricanes moving up the Atlantic seaboard, residents of Georgia's historic city of Savannah are installing inexpensive, yet effective, sea-level monitors that aim, within a year or two, to alert the public to rising coastal sea levels and support a future emergency response system.

These five stories are a sampling of the 18 documented in the comprehensive "How We Respond" project that explores why communities need to act, use science-informed planning and decision-making, and draw from the value of collaboration. The project includes a report and multimedia, and it makes available a host of resources, including research sources such as the 2018 National Climate Assessment and other authoritative studies available on the report's website.

The report, which was released on 16 September, follows AAAS's 2014 "What We Know" report that presented well-established evidence concluding that "human-caused climate change is happening," sending sea levels rising and setting off more frequent and intense weather events, such as heat waves, excessive rainfall, and wildfires.

Communities are devoting time and effort to instituting methodologies to both curb varied impacts of climate change and lower greenhouse gas emissions. The report describes partnerships being forged among scientists, residents, and local and state governments to leverage scientific knowledge, local expertise, and ingenuity to adapt to and mitigate dramatic climate shifts.

Local activities highlighted in the report present a range of innovative tactics to minimize the risks of climate change, including alleviating dangers to human health and to the most vulnerable, such as the elderly and residents in neighborhoods that lack the resources needed to update infrastructure to stave off the forces of climate change.

"We specifically wanted to show how communities across the United States are experiencing different impacts of climate change and are responding in different ways," said Emily Therese Cloyd, director of AAAS's Center for Public Engagement with Science and Technology and a leader of the "How We Respond" project, during a 17 September Facebook Live chat featuring two scientists and a community leader involved in Savannah's Smart Sea Level Sensors initiative. "Another part of this project is to showcase that there are different ways communities might choose to respond and that all those responses are useful and help us make progress."

Some communities, for instance, are devising responses that also preserve economic stability. Consider the Kansas farmers facing depleting groundwater levels that endanger the agricultural industry at the heart of the region's economic livelihood.

Likewise, the efforts by the Oregon hatcheries to curb economic losses due to oyster population declines connected shellfish farmers to scientists able to pinpoint ocean acidification as the factor crippling oyster production levels. The finding helped identify new oyster farming practices. Hatcheries are now aligning the filling of hatchery tanks to times when seawater chemistry is most compatible to oyster hatchlings and the growth of their shells.

PHOTO: IMPACT MEDIA LAB/AAAS

Cooperating across sectors is vital, according to the "How We Respond" report. Key to success, said Cloyd, is "working together, building relationships, opening a dialogue between communities and scientists, and co-creating those responses, learning from one another to find responses that work best for the community."

The story of Savannah's work is rich with multifaceted approaches. The Smart Sea Level Sensor team connected scientists with city planners and public schools, engaging a cross section of residents in the development and placement of sensors around the city and the greater Chatham County coastal area.

The sensors are now beginning to produce local sea-level data. Going forward, long-term flooding patterns are expected to be identified and modeled. As more sensors are installed, additional localized data will lead to the development of the public alert system and help guide future city development and infrastructure projects.

The Smart Sea Level Sensor program was initiated and co-created by two Georgia Institute of Technology professors—Russell Clark, a senior research scientist in the School of Computer Science, and Kim Cobb, an oceanographer and professor of Earth and atmospheric sciences.

Nick Deffley, the sustainability director of Savannah, reached out to Dawud Shabaka, associate director of Savannah's Harambee House, Inc., a community-based education and training organization dedicated to helping the city's underserved communities, to involve the organization in the sensor program.

Community participants are pleased "to collect data in their area and analyze it for use in pushing for public policy changes," said Shabaka, who participated in the Facebook Live chat. He noted that the benefits to Harambee House participants extend beyond educational advantages to the pleasure of being active participants in the project. Harambee House participants, Shabaka noted, are involved in data collection, have access to the online data, and are learning the basics of the data analysis. The project contrasts with other initiatives introduced to the community, he said. "People and companies and so on swoop in and say, 'This is what would be best for you' and then leave. This is not the way this is proceeding," said Shabaka. "Georgia Tech, the city of Savannah, they're working with us."

An early-generation sea-level sensor helps track water levels in Savannah communities.

Beyond engaging residents across communities, local schools are also playing a role, along with city and state officials, scientists, and nonprofit groups. High school engineering students are building and testing a fleet of sensors, made from easily accessible parts. Administrators have developed a curriculum around construction of the sensors that is expected to become a model for other locations along the East Coast that have expressed interest in the program.

Georgia Tech students enrolled in computer design courses also are involved, designing the publicly accessible computer platform for displaying the gathered data and, separately, working on the development of a smart phone app to make the data portable.

"This approach has been very successful," Cobb said. "The students are happy to be working on projects that address meaningful, real-world problems, including local climate change issues."

Cobb envisions the project's framework and the accompanying online Smart Sea Level Sensors platform to be adaptable to a variety of other applications, such as monitoring local air and water quality as well as tracking local shifts in air temperature. "This list goes on and on, and the potential benefits that can be read from a public health perspective," she said. "That's where this project is squarely aimed for the residents of Savannah."

While Clark struck a note of caution in citing the multiple activities that the ongoing project still needs to accomplish before any expansion can be envisioned, he pointed to the project's lasting and most visible reward: "The key benefit for us has been how motivating and inspiring it is in fact to engage the community at this level, and the feedback we get."

Women innovators become STEM ambassadors for girls

AAAS and Lyda Hill Philanthropies seek to raise the profile of women in science through IF/THEN

By **Becky Ham**

Catie Cuan did well in middle school math and science classes but "got made fun of a lot for being smart," she recalled. She was also a dancer and "a gregarious, outgoing person, and I felt like those identities couldn't be reconciled with being a good student."

Today, Cuan is a graduate student in mechanical engineering at Stanford University—and a dancer and choreographer who has used robots in her artistic work. She is also one of the American Association for the Advancement of Science IF/THEN ambassadors, a group of women working in science, technology, engineering, and mathematics who are sharing their stories and serving as high-profile role models for middle school girls.

One of the goals of her ambassadorship, Cuan said, will be to show off the "multiple identities" of women in STEM. "I have a dance background. I'm now working in engineering, and that's OK," she said. "I see this as a distinct opportunity for people to see others who look like them and think, 'Oh, I don't have to be thinking so small about what I can do in the future.'"

The group of 125 ambassadors, announced in September, will serve as part of Lyda Hill Philanthropies' IF/THEN initiative, a $25 million program to fund and elevate women in STEM as role models. For the next 18 months, the ambassadors will participate in a variety of programs, from meeting with local Girl Scout troops to starring in network television series like CBS's *Mission Unstoppable*, a weekly program about women working on cutting-edge STEM projects.

The Lyda Hill Philanthropies approached AAAS to lead the ambassador program in part because of AAAS's extensive success with other STEM education and public engagement initiatives, such as the AAAS Mass Media Science & Engineering Fellowship and the Leshner Leadership Institute for Public Engagement with Science, said Emily Therese Cloyd, director of the AAAS Center for Public Engagement with Science and Technology, who heads up the new ambassador program.

PHOTO: IMPACT MEDIA LAB/AAAS

"The goal of IF/THEN is to shift the way our country—and the world—thinks about women in STEM, and this requires changing the narratives about women STEM professionals and improving their visibility," said Lyda Hill, founder of the philanthropy.

The chance to change that narrative was one of the big reasons Raychelle Burks applied for the ambassadorship. Burks, an analytical chemist and forensic scientist at St. Edward's University in Texas, hopes the program will support girls from historically marginalized communities without telling them they need to "assimilate or conform" to work in the sciences.

"What I want to put out is that we can support girls and women in STEM, that there is a place for them here, and that we will make a place for them, not that they have to conform to preconceived notions and stereotypes," said Burks.

The ambassador program is distinguished by its emphasis on increasing visibility for women in STEM to encompass people who may not consider themselves "scientists," said Cloyd. "We're moving beyond scientists who work at an academic institution and thinking about the ways that a video game designer or a fashion designer might be using STEM every day."

It's an idea that resonated with Ellie Mayfield, who wasn't exposed to many STEM career options as a student. A chance encounter with a woman sport scientist had a profound effect on her future. "The penny only dropped for me that I wanted to be a sport scientist when I saw someone else doing it," she said. Today, she now holds that title with U.S. Soccer and was recently selected to share her story and serve as a AAAS IF/THEN ambassador.

Cuan worked with an industrial robot for her performance piece OUTPUT.

PHOTO: KEVIN BARRY/THOUGHTWORKS ARTS

"I would have liked to know about more options in STEM instead of just being a doctor or an engineer or a veterinarian," agreed Earyn McGee, an ambassador and herpetology graduate student at the University of Arizona. "I didn't realize how broad the category really was."

Other ambassadors, like Charita Castro, want more middle school girls to see how STEM skills can "be deployed in service of causes we care about." Castro leads GoodWeave International, a nonprofit organization dedicated to ending child labor in global supply chains. "I'm a social worker at the core, but I use data and research to inform policies and programs to end child labor, forced labor, and child trafficking," she said.

Middle school girls are already interested in a lot of the technology contained in STEM careers, said Heather Chandler, a multimedia producer who has worked in video game production and now owns an escape room business. "When I tell fifth graders that I had worked on Fortnite, they are just beside themselves," she recalled.

Chandler knows there has been "lots of talk about toxicity" toward women in the gaming industry. "So as an ambassador I want to focus on the positive aspects, to show that it's a multifaceted industry and that we need more women engineers and digital artists and game designers."

"With the group of ambassadors that have been selected, there's a wide range of diversity in terms of geographical location, in terms of ethnicity, in terms of career stage," said Jennifer Carinci, program director of STEM Education Research at AAAS. "It's important educationally for middle school girls to have this exposure."

Carinci, Cloyd, and other AAAS staff led some of the training provided to the ambassadors at an October IF/THEN summit held in Dallas, Texas. The ambassadors learned more about the fundamentals of science communication, developed electronic press kits, worked on strategies to engage with girls on social media and in person, and networked with their fellow ambassadors.

In keeping with their own diversity, the women have a wide variety of plans for their ambassadorships. McGee hopes that the experience will push her further down the path toward someday hosting her own natural history show. Tamar Goulet, a University of Mississippi professor who studies coral reefs, will be Skyping with girls' groups from her Red Sea field site. Burks is in her seventh year of leading DIY Science Zone at GeekGirlCon, a role that she said fits with her pop culture and scientific passions.

"It's hard for me to think of myself as a role model, because I'm a huge goofball," Burks joked. "But I'm the goofball that I get to be, and I'm the nerd that I get to be."

AAAS annual election: Preliminary announcement

The 2019 AAAS election of general and section officers is scheduled to begin later this fall. All members will receive a ballot for election of the president-elect, members of the Board of Directors, and members of the Committee on Nominations. Additionally, members registered in sections (up to three) will receive ballots for the specified section elections. Biographical information for the candidates will be provided along with ballots. The general election slate is listed below. The list of section candidates can be viewed at https://www.aaas.org/governance/2019-election-candidates.

Notice to our international members:
In an effort to conserve resources, AAAS will be sending electronic election ballots to our non–U.S.-based members. In order to ensure that you receive your ballot, please make sure your email is up-to-date with AAAS by logging on to www.aaas.org. 1) Click on "Member Login" (if you have not yet created an account, you will be prompted to do so); 2) After you log in, click on the red "My Profile" button in the upper right-hand corner of the page; 3) Click on "Edit My Contact Information" in the left-hand side bar; 4) Update your email and click on the "Save" button.

If you would like to request a special paper ballot, please send an email with your name and address with your request to elections@aaas.org.

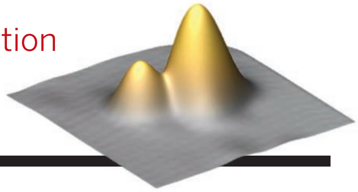
General Election Slate

President-Elect:
Susan Amara, National Institute of Mental Health/NIH
Elsa Reichmanis, Georgia Institute of Technology

Board of Directors:
Cynthia Beall, Case Western Reserve University
Michael J. Feuer, George Washington University
Chad Mirkin, Northwestern University
Alondra Nelson, Social Science Research Center/Institute for Advanced Study

RESEARCH

Single-atom spin manipulation
Yang et al., p. 509



IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

NEURODEVELOPMENT

Close-up of human cerebellar development

Early on, cerebellar development shares similarities across humans, nonhuman primates, and even mice. But differences emerge while development progresses, as cellular and molecular analyses by Haldipur *et al*. now reveal. The rhombic lip persists longer during cerebellar development in humans than in either the mouse or the macaque and generates a pool of neuroprogenitor cells. Similarly, the ventricular zone of the human cerebellum goes a step further than that of the mouse in developing an additional proliferative layer with outer radial glia cells. Transcriptome analysis revealed detailed similarities and differences between progenitor cells of the developing human cerebellum and neocortex. —PJH

Science, this issue p. 454

Fluorescence microscopy image of a developing human cerebellum

INFLUENZA

Alternative influenza target

There is a pressing need for a broadly protective influenza vaccine that can neutralize this constantly varying, deadly virus. Stadlbauer *et al*. turned their attention away from the current vaccine target—the mutable hemagglutinin—and investigated an alternative, less variable virus-coat glycoprotein: neuraminidase. The authors extracted monoclonal antibodies (mAbs) from a human donor naturally infected with the H3N2 virus subtype. In mice, the mAbs were broadly protective against influenza virus A groups 1 and 2 (human, avian, and swine origin) and some influenza B viruses. These mAbs were also therapeutically effective as late as 72 hours after infection. The wide range of reactivity probably relates to the infection history of the donor, whose plasmablasts generated antibodies with long regions that insert into the active site of the neuraminidase enzyme. —CA

Science, this issue p. 499

SOCIAL SCIENCES

Measuring street protest events

Street protests and popular marches are an important form of political expression, and how they are measured shapes our ability to understand their social significance. Fisher *et al*. review the growing research on protest events and crowd forming, with a focus on events following the 2016 U.S. election. They describe best-practice methods for measuring protest size and protester motivation and making such protest data publicly available in real time. Such methods help us understand who protests and why and enable better assessments of the social and political impact of protests. —AC

Sci. Adv. 10.1126/sciadv.aaw5461 (2019).

CELL ENGINEERING

Tea for type 1 and type 2 diabetes

Cell therapy is a promising approach for treating diabetes. Yin *et al*. developed an elegant control system by engineering cells to respond to protocatechuic acid, a metabolite in green tea. Orally ingested protocatechuic acid regulated blood glucose by triggering secretion of insulin or a short variant of human glucagon–like peptide 1 from engineered cells implanted in mouse and

PHOTO: HALDIPUR AND MILLEN

nonhuman primate models of type 1 and type 2 diabetes. This study demonstrates the versatility of synthetic biology for developing remotely controlled cell-based therapies for diabetes. —CC

Sci. Transl. Med. **11**, eaav8826 (2019).

ECONOMICS

Racial bias in health algorithms

The U.S. health care system uses commercial algorithms to guide health decisions. Obermeyer *et al*. find evidence of racial bias in one widely used algorithm, such that Black patients assigned the same level of risk by the algorithm are sicker than White patients (see the Perspective by Benjamin). The authors estimated that this racial bias reduces the number of Black patients identified for extra care by more than half. Bias occurs because the algorithm uses health costs as a proxy for health needs. Less money is spent on Black patients who have the same level of need, and the algorithm thus falsely concludes that Black patients are healthier than equally sick White patients. Reformulating the algorithm so that it no longer uses costs as a proxy for needs eliminates the racial bias in predicting who needs extra care. —TSR

Science, this issue p. 447; see also p. 421

NANOMATERIALS

Milling corundum nanoparticles

High-purity corundum (α-Al_2O_3) nanoparticles could enable applications such as more stable catalyst supports or precursors for high-strength ceramics. Milling of corundum only produces micrometer-scale particles, and direct synthesis from other aluminum oxides that would be likely starting materials, such as γ-Al_2O_3, fails because of the high activation barrier for converting the lattice structure of these cubic close-packed oxides. Amrute *et al*. show that ball milling of boehmite, γ-AlOOH, created ~13-nanometer-diameter corundum nanoparticles of high purity through a mechanically induced dehydration reaction and by the effect of milling impacts on the surface energy of the particles. —PDS

Science, this issue p. 485

BIOGEOGRAPHY

Earthworm distribution in global soils

Earthworms are key components of soil ecological communities, performing vital functions in decomposition and nutrient cycling through ecosystems. Using data from more than 7000 sites, Phillips *et al*. developed global maps of the distribution of earthworm diversity, abundance, and biomass (see the Perspective by Fierer). The patterns differ from those typically found in aboveground taxa; there are peaks of diversity and abundance in the mid-latitude regions and peaks of biomass in the tropics. Climate variables strongly influence these patterns, and changes are likely to have cascading effects on other soil organisms and wider ecosystem functions. —AMS

Science, this issue p. 480; see also p. 425

FERROELECTRICS

Flexible ferroelectrics

High-quality ferroelectric materials, which polarize in response to an electric field, are usually oxides that crack when bent. Dong *et al*. found that high-quality membranes of barium titanate are surprisingly flexible and super-elastic. These films accommodate large strains through dynamic evolution of nanodomains during deformation. This discovery is important for developing more robust flexible devices. —BG

Science, this issue p. 475

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

BEHAVIOR

Dogs' brains and behaviors

Dogs have been bred for traits ranging from herding and hunting to companionship. Attributes such as trainability or highly tuned olfaction characterize different breeds. Using magnetic resonance imaging, Hecht *et al.* scanned the brains of dogs from 10 breed groups defined by behavioral specializations. Although dogs showed diversity in craniofacial shape and brain volume, brain size was not strictly defined by body size. Brain networks were identified that related to behavioral specializations roughly corresponding to social bonding, taste and smell, physical movement, and affective and instinctual functions. The results show that the fingerprints of humans' selective interest in particular types of behaviors is evident in the evolution of dogs' brains. —PJH

J. Neurosci. **39**, 7748 (2019).

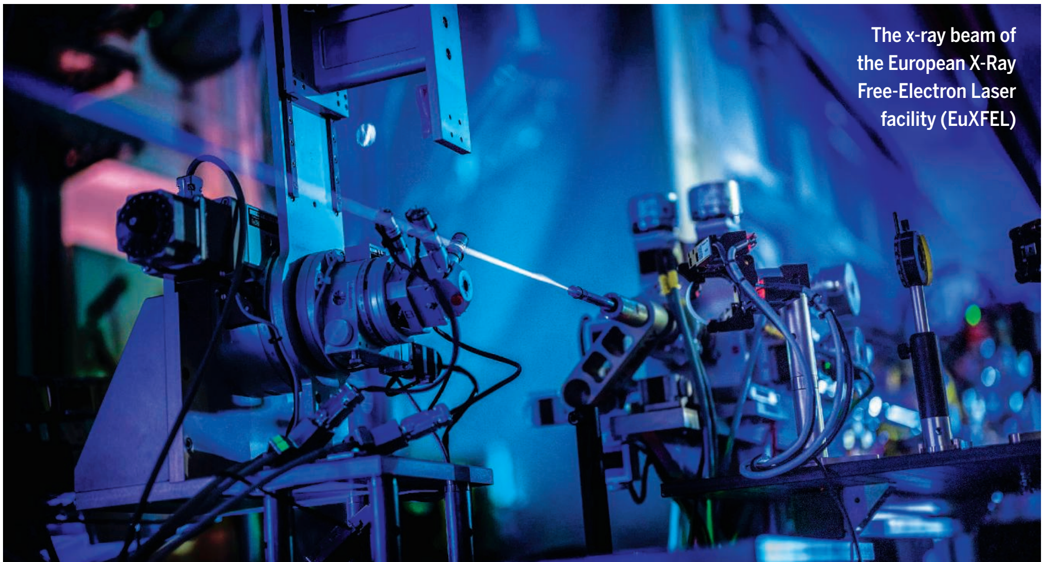
Signs of human selection for behavioral specializations is evident in dogs' brains.

SIGNALING

Network rewiring in cancer

Cancer-causing mutations are likely to modulate existing signaling networks rather than generate newly rewired networks. During tumor growth, tyrosine phosphorylation offers a candidate mechanism for transient alteration of physiological function because of its pivotal role in cell signaling. Working with mouse lung tissue samples, Lundby *et al*. used advances in mass spectrometry and proteomics to examine how oncogenic mutations in the epidermal growth factor

PHOTO: PM IMAGES/GETTY IMAGES



The x-ray beam of the European X-Ray Free-Electron Laser facility (EuXFEL)

X-RAY IMAGING

Adding dynamics to x-ray microscopy

The development of third-generation x-ray sources, such as synchrotrons and free-electron lasers, over the past two decades has provided light-beam brilliance and coherence that can be used to reveal structural details with exquisite spatial resolution. More recent efforts have focused on developing a dynamic capability that would reveal detailed snapshots of ultrafast structural and electronic processes occurring within solid-state and biological samples. By modulating the x-ray beam at megahertz frequencies, Vagovič *et al*. introduce a technique implemented at the European X-ray Free-Electron Laser (EuXFEL) facility that provides microsecond temporal resolution along with micrometer spatial resolution and a better signal-to-noise ratio in the images. The results illustrate a promising route for the further development of temporally resolved x-ray imaging. —ISO

Optica **6**, 1106 (2019).

receptor lead to changes in signals transmitted through its phosphorylated tyrosine residues. The authors focused on one mutation that alters a site three residues away from a phosphorylated tyrosine. This mutation not only abolished recruitment of the normal complex of proteins that interact with the site but also triggered interaction with different proteins that promoted oncogenic signaling. These findings could reveal new targets for anticancer therapies. —LBR

Cell **179**, 543 (2019).

GENOMICS

The thermal adaptation of the proteome

Climate change is likely to cause new and extreme environmental conditions, and we have little understanding of how species will cope. Caddisfly larvae live wrapped in protective cases of grit or fibers among the stones of fast-flowing freshwater streams. Ebner *et al*. examined the proteome of the northern European caddisfly *Crunoecia irrorata* when raised at different temperatures between 10° and 20°C. This species normally lives in the steady environment of groundwater springs, but this study shows that it appears to have compensatory mechanisms that respond to warming. The authors found that differential expression of proteins with functional annotation in candidate thermoregulatory systems increased as temperature increased. It seems that *C. irrorata* has a higher tolerance to temperature variation than previously thought. —LMZ

Molec. Ecol. 10.1111/mec.15225 (2019).

MICROBIOLOGY

Broadening microbiota transplants

Fecal microbiota transplantation (FMT) has been successful at treating *Clostridium difficile* infections, although recent events show there are major risks of antibiotic resistant pathogen transfer (as noted by a recent U.S. Food and Drug Administration alert). FMT is also being investigated to treat bowel cancers and improve therapeutic responses. The gut is not the only site in which the symbiotic microbiota can promote homeostasis. In particular, the vagina contains a specialized microbiota that can be altered after bacterial infection. Lev-Sagie *et al*. treated five patients with intractable and recurrent bacterial vaginosis with vaginal microbiota transplantation (VMT) from healthy donors. Four of the women underwent long-term remission and reconstitution of a normal microbiota. No adverse effects were seen in this study, indicating that VMT could be tested in randomized clinical trials. —GKA

Nat. Med. **25**, 1500 (2019).

SEDIMENT TRANSPORT

Worldwide river sediment flux

The sediment flux coming out of rivers into the ocean is important for river systems and resource management. Li *et al*. compiled a dataset of sediment fluxes for more than 4000 rivers worldwide. The data provide a picture of sediment and water flux that future research data can be compared against. The authors found that many rivers have lower sea-going sediment flux partly as a result of upstream dams and irrigation practices. Changes in water and sediment fluxes because of human impact and climate alters coastal environment and biogeochemical cycles. —BG

Sci. Bull. 10.1016/j.scib.2019.09.012 (2019).

METABOLOMICS

Parallel profiling of an enzyme family

Uncharacterized gene sequences are annotated based on the function of distant relatives and often have very different functions in reality. Kim *et al*. developed a method based on untargeted metabolomics to fill in the gaps and assign precise function to members of the M20 peptidase family, which generally catalyze amide bond cleavage. Viral transduction of each enzyme in mice and comparative analysis of levels of metabolites in liver tissue recovered known enzyme functions for two family members and assigned a new function, dipeptide cleavage, for one of the orphan enzymes in this family. —MAF

Cell Chem. Biol. 10.1016/j.chembiol.2019.09.009 (2019).

PHOTO: EUROPEAN XFEL

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

RENEWABLE ENERGY

A multifaceted future for wind power

Modern wind turbines already represent a tightly optimized confluence of materials science and aerodynamic engineering. Veers *et al*. review the challenges and opportunities for further expanding this technology, with an emphasis on the need for interdisciplinary collaboration. They highlight the need to better understand atmospheric physics in the regions where taller turbines will operate as well as the materials constraints associated with the scale-up. The mutual interaction of turbine sites with one another and with the evolving features of the overall electricity grid will furthermore necessitate a systems approach to future development. —JSY

Science, this issue p. 443

MICROBIOTA

One world, one health

As people increasingly move to cities, their lifestyles profoundly change. Sonnenburg and Sonnenburg review how the shift of recent generations from rural, outdoor environments to urbanized and industrialized settings has profoundly affected our biology and health. The signals of change are seen most strikingly in the reduction of commensal microbial taxa and loss of their metabolic functions. The extirpation of human commensals is a result of bombardment by new chemicals, foodstuffs, sanitation, and medical practices. For most people, sanitation and readily available food have been beneficial, but have we now reached a tipping point? How do we "conserve" our beneficial symbionts and keep the pathogens at bay? —CA

Science, this issue p. 444

IMMUNOLOGY

Commensals rule the MAITrix

Mucosal-associated invariant T (MAIT) cells play an important role in mucosal homeostasis. MAIT cells recognize microbial small molecules presented by the major histocompatibility complex class Ib molecule MR1. MAIT cells are absent in germ-free mice, and the mechanisms by which microbiota control MAIT cell development are unknown (see the Perspective by Oh and Unutmaz). Legoux *et al*. show that, in mice, development of MAIT cells within the thymus is governed by the bacterial product 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil, which rapidly traffics from the mucosa to the thymus, where it is captured by MR1 and presented to developing MAIT cells. Constantinides *et al*. report that MAIT cell induction only occurs during a limited, early-life window and requires exposure to defined microbes that produce riboflavin derivatives. Continual interactions between MAIT cells and commensals in the skin modulates tissue repair functions. Together, these papers highlight how the microbiota can direct immune cell development and subsequent function at mucosal sites by secreting compounds that act like self-antigens. —STS

Science, this issue p. 494, p. 445; see also p. 419

PLANT BIOLOGY

Plant thirst quenched without water

Drought causes many billions of dollars of annual losses to farmers worldwide. Central to a plant's water use efficiency are signaling pathways regulated by the hormone abscisic acid and its receptors. Vaidya *et al*. screened a pool of candidate small molecules and used structure-guided design to optimize the function of an abscisic acid receptor agonist (see the Perspective by Phillips and Sussman). Application of the agonist protected *Arabidopsis*, wheat, and tomato from underwatering. —PJH

Science, this issue p. 446; see also p. 416

INNATE IMMUNITY

NODs require S-palmitoylation to signal

The compartmentalization of proteins within the cell is essential for their function. The addition of lipid molecules redistributes proteins to the cell surface or to membrane-bound organelles. Working in transgenic mice and in tissue cultured cells, Lu *et al*. found that nucleotide oligomerization domain–like receptors 1 and 2 (NOD1 and NOD2), two proteins responsible for detecting bacterial products, required lipid modifications for their recruitment to the cell membrane and function. The specific modification, palmitoylation at a cysteine thiol, was mediated by the enzyme ZDHHC5. Loss of ZDHHC5 or removal of key modification residues in NOD1 and NOD2 abolished their function, compromising antibacterial responses. Human variants of NOD2 display altered palmitoylation, which could help to explain many inflammatory conditions, such as irritable bowel syndrome. —SMH

Science, this issue p. 460

STRUCTURAL BIOLOGY

Complex regulation

The protein kinase mTORC1 controls cellular growth in response to external signals. In the presence of nutrients, it localizes on the surface of the lysosome, where it is activated. The Raptor domain of mTORC1 binds to a complex comprising the protein Ragulator and a heterodimer of the Rag guanosine triphosphatase, which can adopt four different nucleotide conformations depending on nutrient availability. Rogala *et al*. determined the structure of the Raptor-Rag-Ragulator complex at 3.2-angstrom resolution by cryo–electron microscopy. The structure shows why Raptor binds only to a specific nucleotide conformation of the Rag heterodimer and suggests a model for how mTORC1 would dock onto the lysosomal surface, which is a key step in its activation. —VV

Science, this issue p. 468

MUTATION

Genetic background affects variation

Robustness, or the effect of mutations on fitness, can affect the evolutionary trajectory of a species. By introducing a large number of deleterious mutations into many different genetic backgrounds of yeast, Johnson *et al*. found that, for many mutations, the more fit the background, the larger the deleterious effect of the mutation (see the Perspective by Miller). A more-fit lineage is thus less tolerant to deleterious mutations, whereas less-fit lineages can tolerate more mutations. This observation supports a tendency toward diminishing returns for beneficial mutations, which has been shown to influence patterns of adaptation. —LMZ

Science, this issue p. 490; see also p. 418

STRUCTURAL BIOLOGY

Coupled transport

Cation-chloride cotransporters move chloride and cations across the cell membrane and are important in regulating cell volume and setting the chloride concentration inside the cell. Mutations lead to serious diseases, such as epilepsy. Liu *et al*. present the structure of the human potassium-chloride

Published by AAAS

cotransporter KCC1, as determined by cryo–electron microscopy. Based on the structure, functional studies, and molecular dynamics simulations, they propose an ion transport model. The structure provides a framework for interpreting disease-related mutations in potassium-chloride cotransporters. —VV

Science, this issue p. 505

TUMOR IMMUNOLOGY

Interior tumor views

Previous studies indicate that a high frequency of intratumoral neutrophils is associated with a poor clinical prognosis. Si *et al*. used microscopy and imaging techniques to examine how intratumoral interactions between tumor-associated neutrophils (TANs) and tumor-infiltrating lymphocytes (TILs) can affect TIL function. They localized functional cell subsets, which were then used to identify hotspots of TAN-TIL interactions within tumors. Some of these TANs had a distinct phenotype, and their physical association with TILs reduced antitumor functions of those TILs. —CNF

Sci. Immunol. **4**, eaaw9159 (2019).

IMMUNOMETABOLISM

Metabolic quiescence for B cell maturity

Transitional B cell precursors mature into follicular B cells, which are involved in antibody responses. Farmer *et al*. discovered a metabolic checkpoint in this developmental process. Compared with transitional B cells, mouse and human follicular B cells were metabolically quiescent and had increased activation of the kinase AMPK and increased levels of the cell surface ectoenzymes CD39 and CD73, which generate extracellular adenosine. Transitional human B cells that expressed CD73 or were exposed to an AMPK agonist preferentially acquired a follicular B cell phenotype. —ERW

Sci. Signal. **12**, eaaw5573 (2019).

SPIN PHYSICS

Coherent surface spin manipulation

Spin-based quantum information processing requires coherent spin manipulation. Yang *et al*. demonstrate coherent control of surface titanium and iron atom spins on a magnesium oxide surface with a magnetic scanning tunneling microscope tip. Arbitrary sequences of fast electrical pulses delivered to the top induced large electric fields. These fields drove metal-atom movement, which then modulated the tip-atom exchange interaction to create an oscillating effective magnetic field. Advanced spin-control protocols such as Ramsey fringes and Hahn spin echoes revealed quantum dynamics, such as coherent oscillations in a titanium atom dimer assembled on the surface with the tip. —PDS

Science, this issue p. 509

REVIEW SUMMARY

RENEWABLE ENERGY

Grand challenges in the science of wind energy

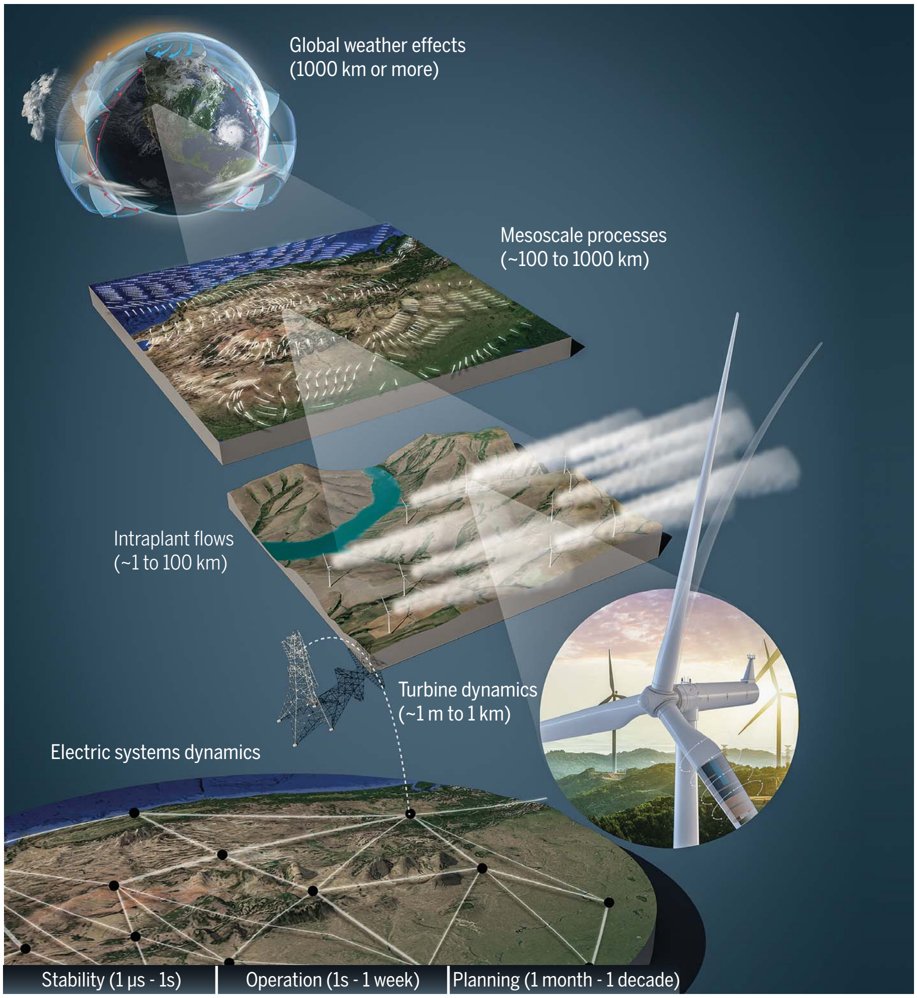
Paul Veers*, Katherine Dykes*, Eric Lantz*, Stephan Barth, Carlo L. Bottasso, Ola Carlson, Andrew Clifton, Johney Green, Peter Green, Hannele Holttinen, Daniel Laird, Ville Lehtomäki, Julie K. Lundquist, James Manwell, Melinda Marquis, Charles Meneveau, Patrick Moriarty, Xabier Munduate, Michael Muskulus, Jonathan Naughton, Lucy Pao, Joshua Paquette, Joachim Peinke, Amy Robertson, Javier Sanz Rodrigo, Anna Maria Sempreviva, J. Charles Smith, Aidan Tuohy, Ryan Wiser

BACKGROUND: A growing global population and an increasing demand for energy services are expected to result in substantially greater deployment of clean energy sources. Wind energy is already playing a role as a mainstream source of electricity, driven by decades of scientific discovery and technology development. Additional research and exploration of design options are needed to drive innovation to meet future demand and functionality. The growing scale and deployment expansion will, however, push the technology into areas of both scientific and engineering uncertainty. This Review explores grand challenges in wind energy research that must be addressed to enable wind energy to supply one-third to one-half, or even more, of the world's electricity needs.

ADVANCES: Drawing from a recent international workshop, we identify three grand challenges in wind energy research that require further progress from the scientific community: (i) improved understanding of the physics of atmospheric flow in the critical zone of wind power plant operation, (ii) materials and system dynamics of individual wind turbines, and (iii) optimization and control of fleets of wind plants comprising hundreds of individual generators working synergistically within the larger electric grid system. These grand challenges are interrelated, so progress in each domain must build on concurrent advances in the other two. Characterizing the wind power plant operating zone in the atmosphere will be essential to designing the next generation of even-larger wind turbines and achieving dynamic control of the machines. Enhanced forecasting of the nature of the atmospheric inflow will subsequently enable control of the plant in the manner necessary for grid support. These wind energy science challenges bridge previously separable geospatial and temporal scales that extend from the physics of the atmosphere to flexible aeroelastic and mechanical systems more than 200 m in diameter and, ultimately, to the electrical integration with and support for a continent-sized grid system.

OUTLOOK: Meeting the grand research challenges in wind energy science will enable the wind power plant of the future to supply many of the anticipated electricity system needs at a low cost. The interdependence of the grand challenges requires expansion of integrated and cross-disciplinary research efforts. Methods for handling and streamlining exchange of vast quantities of information across many disciplines (both experimental and computational) will also be crucial to enabling successful integrated research. Moreover, research in fields related to computational and data science will support the research community in seeking to further integrate models and data across scales and disciplines. ■

ON OUR WEBSITE
Read the full article at http://dx.doi.org/10.1126/science.aau2027



The cascade of scales underlying wind energy scientific grand challenges. Length scales from weather systems at a global level down the boundary layer of a wind turbine airfoil and time scales from seasonal fluctuations in weather to subsecond dynamic control and balancing of electrical generation and demand must be understood and managed.

ILLUSTRATION: JOSH BAUER AND BESIKI KAZAISHVILI, NREL

The list of author affiliations is available in the full article online.
*Corresponding author. Email: paul.veers@nrel.gov (P.V.); kady@dtu.dk (K.D.); eric.lantz@nrel.gov (E.L.)
Cite this article as P. Veers *et al.*, *Science* **366**, eaau2027 (2019). DOI: 10.1126/science.aau2027

REVIEW

RENEWABLE ENERGY

Grand challenges in the science of wind energy

Paul Veers[1]*, Katherine Dykes[2]*, Eric Lantz[1]*, Stephan Barth[3], Carlo L. Bottasso[4], Ola Carlson[5], Andrew Clifton[6], Johney Green[1], Peter Green[1], Hannele Holttinen[7], Daniel Laird[1], Ville Lehtomäki[8], Julie K. Lundquist[1,9], James Manwell[10], Melinda Marquis[11], Charles Meneveau[12], Patrick Moriarty[1], Xabier Munduate[13], Michael Muskulus[14], Jonathan Naughton[15], Lucy Pao[16], Joshua Paquette[17], Joachim Peinke[3,18], Amy Robertson[1], Javier Sanz Rodrigo[13], Anna Maria Sempreviva[2], J. Charles Smith[19], Aidan Tuohy[20], Ryan Wiser[21]

Harvested by advanced technical systems honed over decades of research and development, wind energy has become a mainstream energy resource. However, continued innovation is needed to realize the potential of wind to serve the global demand for clean energy. Here, we outline three interdependent, cross-disciplinary grand challenges underpinning this research endeavor. The first is the need for a deeper understanding of the physics of atmospheric flow in the critical zone of plant operation. The second involves science and engineering of the largest dynamic, rotating machines in the world. The third encompasses optimization and control of fleets of wind plants working synergistically within the electricity grid. Addressing these challenges could enable wind power to provide as much as half of our global electricity needs and perhaps beyond.

Abundant, affordable energy in many forms has enabled notable human achievements, including modern food and transportation infrastructure. Broad-based access to affordable and clean energy will be critical to future human achievements and an elevated global standard of living. However, by 2050, the global population will reach an estimated 9.8 billion, up from ~7.6 billion in 2017 (*1*). Moreover, Bloomberg New Energy Finance (BNEF) estimates suggest that annual global electricity demand could exceed 38,000 terawatt-hours per year by 2050, up from ~25,000 terawatt-hours in 2017 (*2*). The demand for low- or no-carbon technologies for electricity is increasing, as is the need for electrifying other energy sectors, such as heating and cooling and transport (*2*–*4*). As a result of these two partially coupled megatrends, additional sources of low-cost, clean energy are experiencing increasing demand around the globe. With a broadly available resource and zero-cost fuel, as well as exceptionally low life-cycle pollutant emissions, wind energy has the potential to be a primary contributor to the growing clean energy needs of the global community.

During the past decade, the cost of three major electricity sources—wind power, solar power, and natural gas—has decreased substantially. Wind and solar are attractive because their low life-cycle emissions offer public health and broader environmental benefits. Leading energy forecasters such as consultancies, nongovernmental organizations, and major energy companies—and specifically BNEF, DNV GL, the International Energy Agency (IEA), and BP—anticipate continued price parity among all of these sources, which will likely result in combined wind and solar supplying between one- and two-thirds of the total electricity demand and wind-only shares accounting for one-quarter to one-third across the globe by 2050 (*3*–*6*). Tapping the potential terawatts of wind energy that could drive the economic realization of these forecasts and subsequently moving from hundreds of terawatt-hours per year to petawatt-hours per year from wind and solar resources could provide an array of further economic and environmental benefits to both local and global communities.

From a business perspective, at just over 51 gigawatts of new wind installations in 2018 (*7*) and more than half a terawatt of operating capacity, the global investment in wind energy is now ~$100 billion (U.S. dollars) per annum. The energy consultant DNV GL predicts that wind energy demand and the scale of deployment will grow by a factor of 10 by 2050, bringing the industry to the trillion-dollar scale (*6*) and positioning wind as one of the primary sources of the world's electricity generation.

However, to remain economically attractive for investors and consumers, the cost of energy from wind must continue to decrease (*8*, *9*). Moreover, as deployment of variable-output wind and solar generation infrastructure increases, new challenges surface related to the adequacy of generation capacity on a long-term basis and short-term balancing of the systems—both of which are critical to maintaining future grid system stability and reliability (*10*–*12*).

A future in which wind energy contributes one-third to more than one-half of consumed electricity, and in which local levels of wind-derived power may exceed 100% of local demand, will require a paradigm shift in how we think about, develop, and manage the electric grid system (*10*–*14*). The associated transformation of the power system in high-renewables scenarios will require simultaneous management of large quantities of weather-driven, variable-output generation as well as evolving and dynamic consumption patterns.

A key aspect of this future system is the availability of large quantities of near-zero marginal cost energy, albeit with uncertain timing. With abundant near-zero marginal cost energy, more flexibility in the overall electricity system will allow many different end users to access these "cheap" energy resources. Potential use cases for this energy could entail charging a large number of electric vehicles, providing inexpensive storage at different system sizes (consumer to industrial) and time scales (days to months), or channeling into chemicals or other manufactured products (sometimes referred to as "power-to-X" applications).

A second key aspect of this future system is the transition from an electric grid system centered on traditional synchronous generation power plants to one that is converter dominated (*15*). This latter paradigm reduces the physical inertia in the system currently provided by traditional power plants while increasing reliance on information and digital signals to maintain the robustness and power quality of the modern grid (*12*).

Historical development of wind energy science

Wind power was harnessed early in the history of civilization, first to propel sailing vessels and later to drive windmills that were often used for grinding grain and pumping water.

[1]National Renewable Energy Laboratory (NREL), Golden, CO, USA. [2]Department of Wind Energy, Technical University of Denmark, Kongens Lyngby, Denmark. [3]ForWind - Center for Wind Energy Research, Oldenburg, Germany. [4]Wind Energy Institute, Technical University of Munich, Garching, Germany. [5]Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. [6]WindForS - Wind Energy Research Cluster, Stuttgart, Germany. [7]Recognis Oy, Espoo, Finland. [8]Kjeller Vindteknikk Oy, Espoo, Finland. [9]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Boulder, CO, USA. [10]Department of Mechanical and Industrial Engineering, University of Massachusetts Amherst, Amherst, MA, USA. [11]NOAA Global Systems Division, Boulder, CO, USA. [12]Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD, USA. [13]National Renewable Energy Center of Spain, Navarre, Spain. [14]Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, Trondheim, Norway. [15]Department of Mechanical Engineering, University of Wyoming, Laramie, WY, USA. [16]Department of Electrical, Computer and Energy Engineering, University of Colorado, Boulder, Boulder, CO, USA. [17]Sandia National Laboratories, Albuquerque, NM, USA. [18]Institute of Physics, University of Oldenburg, Oldenburg, Germany. [19]Energy Systems Integration Group, Reston, VA, USA. [20]Electric Power Research Institute, Palo Alto, CA, USA. [21]Lawrence Berkeley National Laboratory, Berkeley, CA, USA.
*Corresponding author. Email: paul.veers@nrel.gov (P.V.); kady@dtu.dk (K.D.); eric.lantz@nrel.gov (E.L.)

However, it was not until the early 20th century, thanks to the pioneering work of Albert Betz and others in the burgeoning field of aerodynamics, that a foundation for wind energy science was developed (*16*) and specifically applied to electricity generation. Leveraging design principles informed by the science, "wind dynamos" were produced and deployed globally to provide power to those who could not yet access the larger electricity grid. As the modern electric system grew worldwide, however, it was the oil crisis of the 1970s that rekindled interest in renewable energy technologies and led to commercial adoption of grid-integrated wind energy systems.

Since that time, wind energy has grown from a niche resource to supply ~5% of global electricity generation (*7*). Levels in some countries have extended well beyond this global average, reaching 10%, 20%, or more in several countries around the world (*17*). This growth in wind energy deployment was associated with a marked decline in the levelized cost of energy (LCOE) driven by both research and technological learning curves (*18*). Because of the nearly half-century of sustained innovation in wind energy, levelized costs are now a fraction of the early-1970s costs. Currently, costs for wind energy are ~\$0.04/kilowatt-hour (*9*, *17*) and are competitive, without subsidies, with other newly installed sources of electricity generation in a growing number of regions (*19*, *20*). The reduction in LCOE over recent decades has spurred further deployment of wind energy with annual global installations reaching >50 gigawatts and cumulative operating capacity of wind energy of more than half a terawatt (see Fig. 1).

Three fundamental drivers have reduced the cost of wind energy to date: increased hub height, power rating, and rotor diameter. These can be understood using the fundamental equation for wind turbine energy capture

$$P = \frac{1}{2}\rho C_p A V^3$$

where P is the instantaneous power produced, ρ is the air density, C_p is the power coefficient (or overall machine aerodynamic-mechanical-electrical performance measure), A is the swept area of the rotor, and V is the free-stream air velocity. The design of the machine affects access to higher V, as well as performance, C_p, and A. Increasing hub height reduces the influence of the surface friction, allowing wind turbines to operate in higher-quality resource regimes where wind velocities are higher, with a compounding effect on power production. Larger generator capacity coupled with power electronics—which enable variable-speed operation—provides more power produced per machine installed at a given location (assuming a constant C_p). More power per turbine allows fewer turbine installations, lower balance-of-system costs, and fewer moving parts (for a given level of power capacity), thereby enhancing reliability. In addition, variable speed with constant frequency output allows the turbine to operate at peak C_p across a wide range of wind speeds for increased energy capture. The third fundamental driver is larger, more efficiently designed wind turbine rotors that sweep a greater area with advanced blades using less material. Larger rotors capture more of the energy passing by each turbine, and because blade lengths can be increased while many other costs remain fixed, they provide a substantial cost reduction on a dollar-per-unit energy basis. In addition, as the size of the rotor grows relative to the generator rating, the turbine will have a lower rated wind speed and operate more frequently at full power output. Although today's optimized, low-cost, and reliable machines—with hub heights at 100 m or more, blade lengths reaching well beyond 50 m, and power ratings of 5 megawatts and up—are the beneficiaries of decades of fundamental research and innovation, the next generation of improvement will depend on further advancements in knowledge and technology.

In this context, continued wind technology innovation is challenging, partly because of classical problems. For example, simply scaling the machine rotor diameter and rated power runs afoul of the "square–cube law," as it is commonly known within the wind industry and research community. Assuming a constant wind speed across the rotor plane, the amount

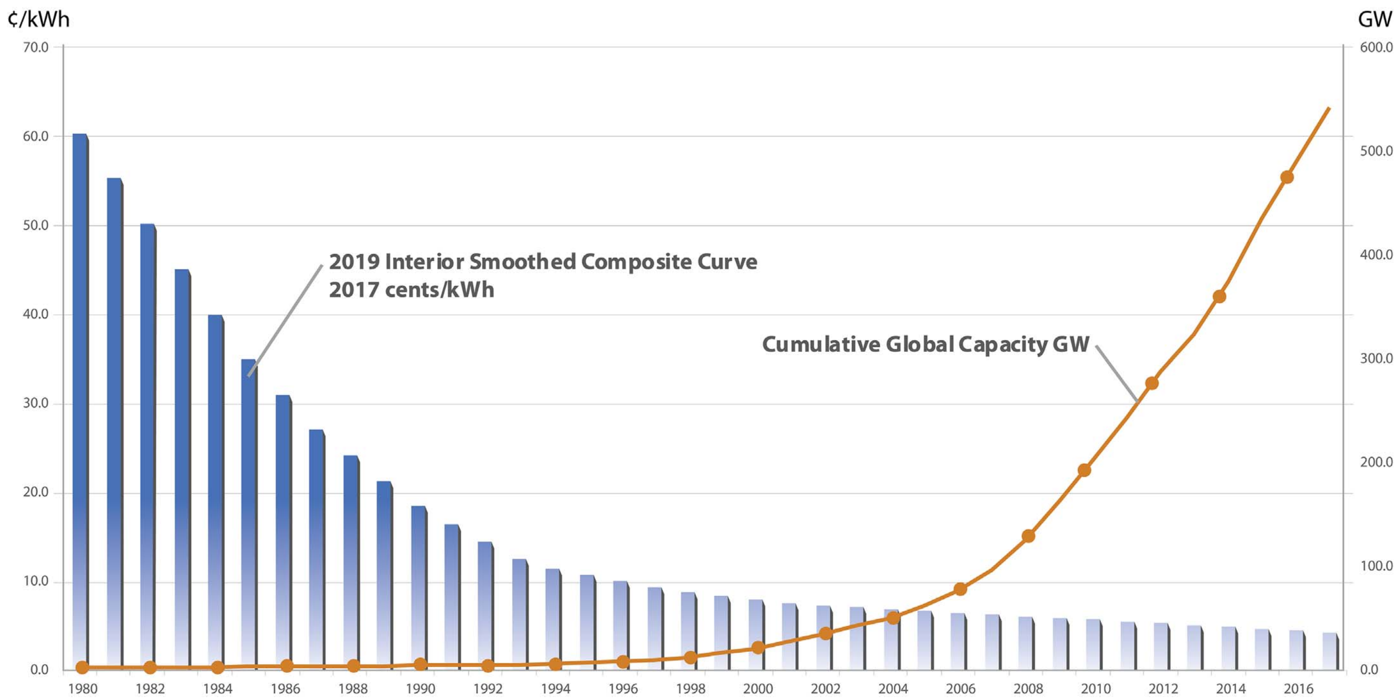


Fig. 1. Global cumulative installed capacity (in gigawatts) for wind energy and estimated LCOE for the U.S. interior region in cents per kilowatt-hour from 1980 to the present. Historical LCOE data are from (*17*) and (*20*) and have been verified for all but 5 years with the U.S. wind industry statistics database detailed in (*17*). LCOE data have been smoothed with a combination of polynomial best fit and linear interpolations to emphasize the long-term trends in wind energy costs. Historical installed capacity data are from the database detailed in (*17*), the Global Wind Energy Council, and the American Wind Energy Association.

of incorporated material scales with volume (the cube), whereas the energy capture scales only with the area of the rotor (the square). Although economies in the balance-of-system costs and elsewhere in the system mitigate the impacts of this particular problem, integrated innovation in all aspects of wind turbine design is necessary to achieve meaningful gains in per-unit energy costs.

Future wind technology innovation is further challenged by the extent of progress that has been achieved already and can be illustrated by focusing on the wind turbine blade. A modern blade is far more sophisticated in aerodynamic design, use of materials, manufacturing process, and structure than ever before (*21*, *22*) and has fundamentally different features than other aerodynamic applications such as airplane wings. Figure 2 shows a comparison of the design features of a current state-of-the-art blade versus a blade from the 1980s. Some key innovations include higher tip speeds to reduce torque and minimize drivetrain weight; higher speed and high-lift airfoils for a more slender, lighter blade; and innovative tip shapes to mitigate noise. Innovations over time have led to modern blades that are 90% lighter than the 1980s blade would be if simply scaled to current lengths. Examples include aeroelastic tailoring, which passively reduces the loads through coupling blade bending and twist; thicker flat-back airfoils, which enable improved aerodynamic performance from the load-bearing section near the hub; add-ons such as vortex generators and flow fences; and a variety of manufacturing improvements (*23*, *24*).

Grand challenges in wind energy research

The research challenges that are critical to realizing the full potential of wind energy stem from the complex and highly coupled phenomena that cross many physical and temporal scales relevant to wind energy and the broader power system. To extract maximum value at minimum cost while maintaining power system reliability and resiliency, it is important to look from global weather phenomena to regional weather activity to complex local flows, and ultimately, to the responses of the turbines within the power plant (Fig. 3). At the same time, the behavior of the wind resource varies greatly by location, as the wind resource behaves differently offshore, across plains, and over mountains. Moreover, a fleet of wind power plants must be in sync with the demands of power system operators as well as consumers at time scales ranging from the subsecond to the decade.

Although the European Academy of Wind Energy envisioned a comprehensive agenda for research challenges in wind energy in 2016 (*25*), the scale of further technology advancement and the magnitude of the challenge associated with relying on wind energy for one-third or more of the global electricity demand necessitated further examination of research needs. This additional effort sought to sharpen the focus of the wind energy research community and identify critical skills and capabilities from the broader scientific and research community that will be necessary to enable use of wind energy at very high levels. To address this need, a group of international wind power experts came together in a series of IEA Wind Technology Collaboration Programme meetings beginning in October 2017 to explore and articulate innovation pathways and associated research challenges that, if addressed, would position wind energy as a primary supplier of the world's electricity needs at levels of one-third to one-half or even more [see (*26*) for detailed findings]. These challenges were then synthesized into a set of three grand challenges requiring a comprehensive and integrated research program across many scientific disciplines (*27*).

First grand challenge: Improved understanding of atmospheric and wind power plant flow physics

Wind energy ensues from the uneven heating of Earth's surface and the Coriolis forces of Earth's rotation. It is a heterogeneous resource highly dependent on geographic location and local terrain, whether mountainous or relatively flat, in plains or deserts. The wind resource over the ocean depends on a different set of meteorological drivers, including sea and land breezes, proximity to land, water versus air temperature, and wave height. Even in specific locales, the wind varies between day and night and across seasons. Wind turbines reside in the lower levels (e.g., <300 m) of the atmospheric or planetary boundary layer. This region is referred to as the surface layer and is where obstructions such as trees, buildings, hills, and valleys cause turbulence and reduce the speed of the wind. Because the sources of wind originate in global meteorological phenomena and the subsequent extraction of energy from the wind occurs in the surface layer, the scales and physics involved reach further than those of many, if not all, other large-scale dynamic systems. Historically, simplification of the overall physics associated with different scales allowed narrowly focused research communities to thrive independently. In this context, wind designers have avoided the need to model large-scale weather effects by focusing on the flow over short durations and affected only by local topography. This approach requires assumptions such as stationarity (consistency over time) and surface-layer similarity (where momentum and heat fluxes are uniform with height) and separates the physics into flows at large mesoscales versus plant-level microscales (*28*–*30*).

More specifically, the mesoscale and the microscale are numerically modeled in fundamentally different ways, thereby making the assessment of atmospheric effects on wind plants that span these scales extremely difficult. The mesoscale processes, which influence local weather, are on the order of 5 to hundreds of kilometers in size and are typically modeled using grid spacing of 1 to 10 km. Microscale processes, the phenomena that drive wind turbine and plant behavior, extend well below 1 km and have grid spacing between 5 and 100 m horizontally. Vertically, microscale model resolution may go to within a few meters of the surface, but the flow is treated as an average over the horizontal grid spacing, making resolution of flow details that affect a wind turbine impossible. If the length scale of the process is much greater

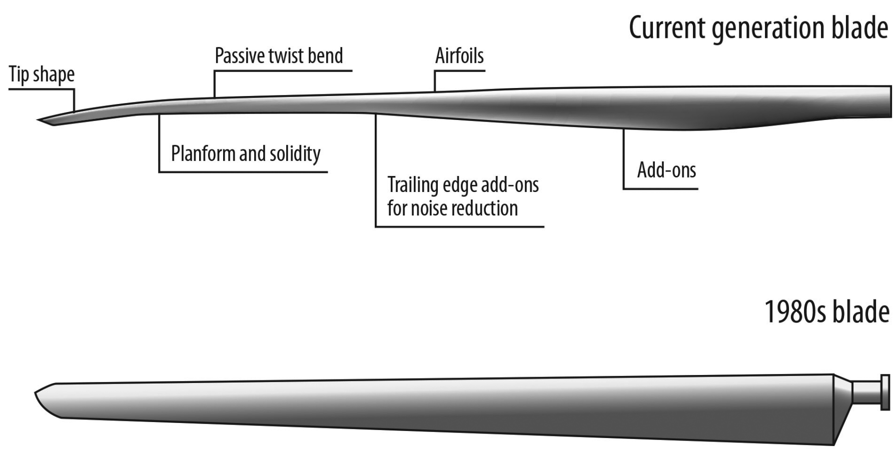


Fig. 2. Wind turbine blade innovation comparing a modern commercial blade (top) and a commercial blade from the mid-1980s (bottom) scaled to the same length. The modern blade is 90% lighter than the scaled 1980s technology.

NATIONAL RENEWABLE ENERGY LABORATORY (NREL) BASED ON A CONCEPT BY HENRIK STIESDAL AND KENNETH THOMSEN (SIEMENS GAMESA)

than the model grid spacing, the process is explicitly resolved; if the length scale of the process is much less than the model grid spacing, the process is parameterized or simplified.

Atmospheric phenomena that span approximately 1.5 to 0.5 km exist at the interface of mesoscale and microscale processes (Fig. 3). This zone, dubbed the "terra incognita" (unknown territory) by Wyngaard (*31*), spans atmospheric processes and their respective physical models of fundamentally different character and understanding. At spatial scales greater than 1.5 to 0.5 km, models resolve only average flows, parameterizing the effects of turbulence implicitly, whereas models over smaller distances resolve turbulence explicitly and simulate the time-varying, stochastic flow fields. Linking the two depends on a comprehensive understanding of the nature of the transition, an understanding that is currently elusive (*32*, *33*).

The scale that characterizes the terra incognita has become increasingly important as the economics associated with wind turbines and plants have pushed blade tip heights and rotor sizes to 200 m, with expectations for even larger sizes in the future. At this scale, wind turbines are affected by turbulent flow features that are driven by mesoscale phenomena and play out within the terra incognita. Specifically, the spatial scale of these atmospheric processes begins to match the scale and height of the turbine rotor, and accordingly, the physics of this poorly understood zone becomes critical to ensuring optimal design and performance of individual turbines and entire wind power plants (*34*, *35*).

Closely associated and interlinked with the mesoscale-to-microscale transition are additional challenges in understanding the flow physics of wind power plants. First, flow propagating through the wind power plant depends on microscale flow effects from the combined influence of the atmosphere and terrain on land, the sea surface offshore, or both. Second, interaction with the turbines themselves modifies the flow as it passes through each subsequent row of turbines in the plant.

Although past use of simplified physical models and basic observational technology has allowed for installation of wind power plants and predictions of performance in a variety of terrain types, there are still major gaps in our knowledge about wind flows in complex terrain or under varying atmospheric stability conditions that can change over the course of a day or season (*34*, *36*). Moving to offshore wind power introduces additional coupled physics of the meteorological-oceanographic (i.e., the "metocean") environment, where a nontrivial modeling uncertainty remains, especially with breaking or irregular waves, atmospheric stability, and tropical storms (*37*).

The creation of wakes—low-energy regions in the flow caused by extracting energy from that flow—is illustrated in Fig. 3 as haze streaming behind the turbines in the microscale flow graphic and behind the full wind plants in the mesoscale. The existence of wakes further complicates the process of understanding both the overall plant performance (energy production) and the loads experienced by the turbines (translating to capital and operational costs). Wind turbine wakes are complex: Their behavior varies with turbine size and design as well as different inflow and turbine operating conditions and may have long-lasting effects, both within a given wind plant and between neighboring plants (*33*, *38*–*40*).

The impact of the wake of one power plant on downstream plants and the local environment has also been explored with mesoscale modeling tools (*41*–*43*) as well as in situ measurements (*35*, *44*–*47*) but is not yet well understood. Measurable changes in the local microclimate can influence surface temperature, humidity, and agriculture (*35*, *44*), but these effects are also highly variable and difficult to predict. This is even true of offshore wind farm microclimates (*47*). Some investigators question at what point regional development of wind reaches saturation and then diminishing returns (*48*, *49*), but opinions vary widely. Wind farm wakes also change with atmospheric stability, which complicates the ability to assess interference (*50*, *51*). Finally, the regional intensity of the resource may be affected by changes in the climate (*52*), raising issues of siting and profitability for future wind power plant development. For more detailed research questions specific to relevant subdisciplines, including meteorology research and fluid turbulence, see (*53*) and (*54*), respectively.

Recent advances in measurement technologies for remote sensing (using lasers, acoustics, or radar to measure atmospheric phenomena) are being used to characterize wakes as they form and propagate through wind power plants (*55*–*60*). However, additional advances in such technologies and their use in measurement campaigns in a wide range of environmental conditions are needed to further resolve the physics of wakes and their impact on the individual turbine, overall plant, and interplant operation. In the offshore metocean environment, it is even more challenging to collect measurements (*61*). In these cases, Sempreviva *et al.* show how integration of data from lighthouses, ships, and buoys can be incorporated with remote sensing and

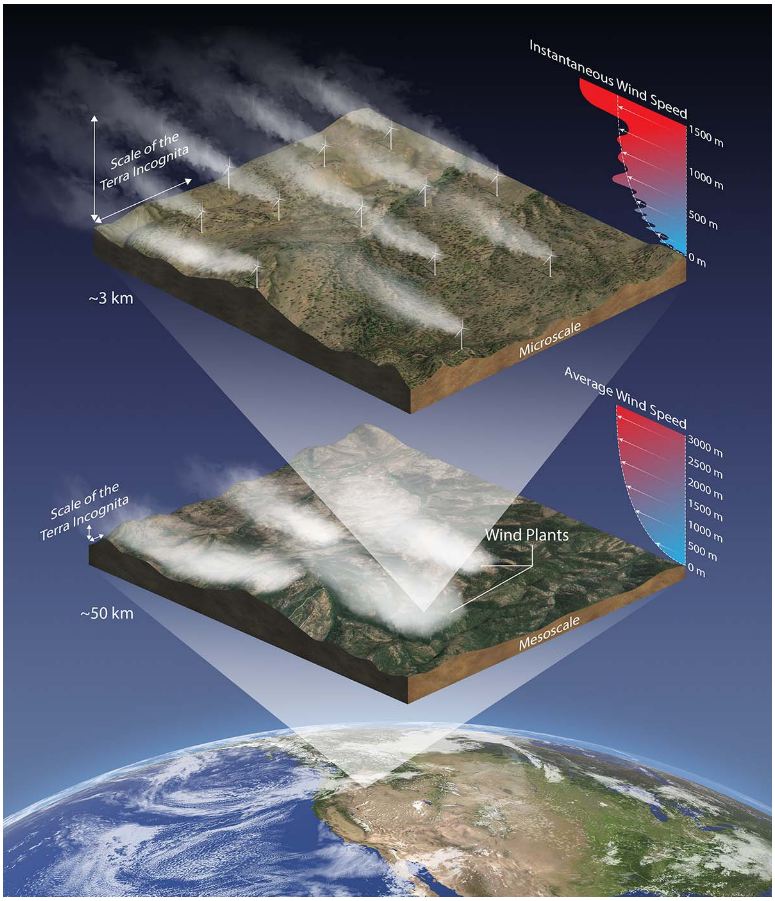


Fig. 3. Relevant wind power scales across space—from large-scale atmospheric effects in local weather at the mesoscale to inter- and intraplant flows and topography at the microscale.

ILLUSTRATION: BESIKI KAZAISHVILI, NREL

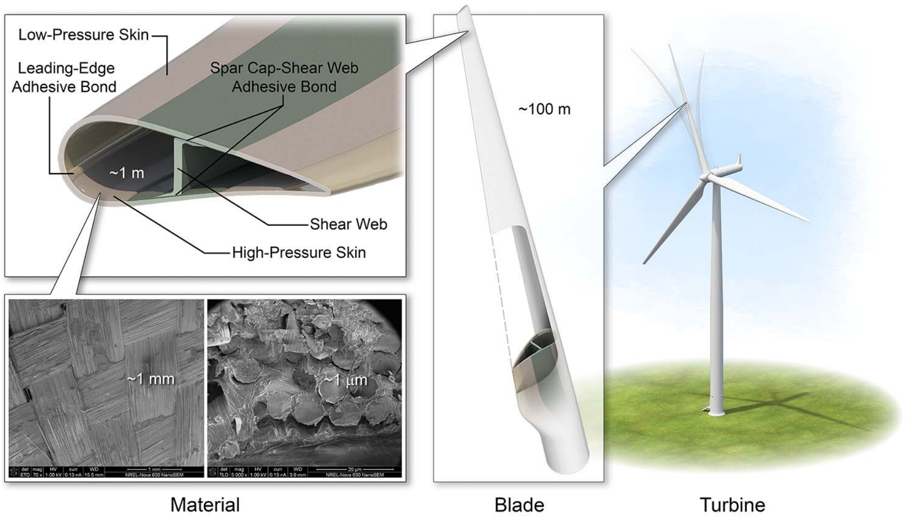


Fig. 4. Wind turbine blades are complex composite shell structures in which small-scale manufacturing flaws can grow because of the incessant turbulence-driven loading that can cause large-scale problems.

modeling (62). Such platforms can extend the reach of measurements but impose their own limitations, illustrating the need for greater innovation in instrumentation and techniques.

Second grand challenge: Aerodynamics, structural dynamics, and offshore wind hydrodynamics of enlarged wind turbines

An operating wind turbine may appear to be very still, apart from the rotation of the blades, yet the entire system is constantly flexing because of forces and moments in all directions and over its entire operating life of 20 years or more. Underpinning this constant movement are important couplings between the wind flow into and through a plant and the turbine responses and interaction with that flow. The dynamics of the turbine response over its lifetime requires meaningful further research.

In the past several decades, numerical wind turbine simulation capabilities that incorporate state-of-the-art knowledge of wind turbine physics (e.g., coupling aerodynamics, structural dynamics, control systems, and even hydrodynamics for offshore applications) have enabled the wind industry to design machines that deliver efficient power for years on end, surviving all weather extremes. As a result, wind turbines have grown to become the largest flexible, rotating machines in the world—massive civil engineering structures that must operate continuously for 20 years or more (a typical design and financial amortization period) under constant complex loading. Blade lengths are approaching 80 m and towers are growing well above 100 m for maximum tip heights, often exceeding 200 m, equivalent to a building more than 60 stories high. To put these dimensions in another context, three of the largest passenger aircraft, an Airbus A380-800s with a wingspan of 80 m, could fit within the swept area of one wind turbine rotor.

However, for both land-based and offshore applications, the industry is seeking even larger turbines that access higher wind speeds aloft and provide further economies of scale, reducing manufacturing, installation, and operational costs per unit of plant capacity. As machines continue to increase in size, several important research questions pertaining to wind turbine dynamics must be addressed. These questions involve the interaction of turbine dynamics with the atmosphere, wakes, and other sources of complex inflow to the rotor, as well as the high Reynolds number and aeroelastic behavior of very large and flexible machines. In addition, the dynamics associated with deployment offshore in conditions such as extreme weather events or deployment on floating platforms with additional degrees of freedom in movement must also be explored.

The larger turbines of the future would operate partly above the often-studied atmospheric surface layer where they could encounter substantial variation in inflow because of poorly characterized factors, such as shear (vertical differences in wind speed), veer (vertical differences in wind direction), and wakes of upstream turbines. The challenge lies not only in understanding the atmosphere but in deciphering which factors are critical in both power-generation efficiency and structural safety. The design perspective must increasingly consider the interdependence of the meso-to-microscale transition and the turbine dynamics to assess, accurately predict, and manage loads (33, 37, 63, 64).

The aerodynamic assumptions themselves are increasingly being questioned. The interaction between a highly variable inflow and the unsteady aerodynamics of the moving and deforming blades is pushing the limits of current theory. Recent experiments at the largest scales now possible by the Danish Technical University (65) suggest that the interaction of these large blades with turbulence of different intensities could be affecting the fundamental lift and drag characteristics of the airfoil, which is not a consideration at smaller scales (66). Because experimental ground truth is difficult to obtain in the uncontrollable atmosphere, researchers are looking to the next generation of exascale supercomputers to provide insight that bridges the blade surface boundary layer (in micrometers) to the planetary boundary layer (in kilometers) (67, 68).

The elastic displacements of these highly flexible structures complicate the aerodynamics, creating complex aeroelastic behavior of the machines as they grow in size. Blades moving through air shed vorticity, which is normally convected downstream and away from a relatively stiff structure. When the blades flex into and out of the wind, the rotor interacts with its own vorticity, calling the accuracy of the design assumptions into question. Additionally, structural dynamics of blades incorporating composite materials, built-in curvature and sweep, and large nonlinear deflection (including torsion and bend-twist coupling) further complicate models of the physics (69) and the assessment of crucial design aspects such as stability (70, 71). In fact, although aeroelastic stability has typically not been a key design driver for rotor blades up to now, the situation may change for future highly flexible and large rotors. Indeed, stability analysis is necessary for avoiding resonance phenomena, ensuring a safe margin to flutter, and understanding the effects of low damped modes on vibrations and loading.

Offshore wind power plants require the combined modeling of aerodynamics and the hydrodynamic forces from waves and currents. Although offshore structures for a variety of applications (including oil drilling) have been designed and constructed for decades, the aerodynamic and hydrodynamic forces have not been of similar magnitudes, nor have they interacted to such an extent that coupled analysis was required (72–74). To explore configurations for offshore support structures specific to wind energy, the hydrodynamic models will need to include the combined nonlinearity and irregularity of sea states, breaking waves (75), viscous effects on bluff bodies at high Reynolds numbers, vortex-induced vibrations, dynamic soil-structure interactions of the seabed foundation, and more (73, 76, 77). Particularly relevant for these offshore applications are the extreme weather conditions, such as hurricanes or tropical cyclones, that are prevalent in many areas of the world where offshore

wind energy deployments are planned, such as on the East Coast of the United States or in the Pacific Ocean near Korea, Taiwan, and Japan (*78*, *79*). Han *et al.* outline the factors that must be taken into consideration when building an offshore wind power plant in regions affected by hurricanes (*80*).

Floating offshore systems, which promise to enable wind energy in large areas of the ocean with water depths of ~60 m or more, have additional degrees of freedom in the motion of the turbine platform (*74*). The uncertainty associated with the rotor interacting with its own vorticity for very large blades is amplified if the entire rotor is rocking into and out of its own wake (*81*), as could happen on a floating foundation (*82*). This aerodynamic problem is compounded by hydrodynamic complexity because the large motions undergone by these turbines violate hydrodynamic theory assumptions typically used in marine structural design (*74*, *83*). The coupled stability analysis of such complex aero-hydro-servo-elastic systems is a problem that has not been thoroughly studied in the past.

New materials and manufacturing methods are an integral part of enabling the development of these structures. Understanding the dynamics will help establish the design requirements, but materials and manufacturing breakthroughs will be needed to enable low-cost, reliable machine designs. Although wind energy has benefited from materials innovation in the past several decades—through fiber-reinforced composites, rare-earth magnets, semiconductors for power electronics, lubricants, and more—there is still a critical need to improve materials performance for particularly difficult environmental conditions and operational loads. The specific challenges related to materials science and engineering for wind energy are the need for materials to have tuned or customized properties for the specific application, as well as the need to be commoditized—that is, easily mass produced at very low cost. Ready recyclability is another desirable attribute (the blade shown in Fig. 4 is one example of a difficult-to-recycle component). The turbine blade and various subcomponents must be integrated at large scales (1 to ≥100 m), but their properties need to be tailored at small scales (1 μm to ≥1 mm).

The blade requires sufficient stiffness to avoid striking the tower, flexibility to adapt continuously to changing wind conditions, durability to last for two decades, and a surface that fights erosion while shedding moisture and dirt—all at commodity prices. Modern blades still use materials similar to those of the 1990s machines, which were based on low-cost composite fibers and durable epoxy resins. Innovations in the resin matrix, fiber reinforcement, and core materials, as well as adhesives and manufacturing protocols, are needed to achieve improved strength, stiffness, and weight properties at very low cost. Blade manufacturing would be markedly improved if thermoplastic resins could be proven viable for blades, allowing secondary welding of the composite structural elements and, perhaps most importantly, recyclability at the end of life (*84*). Beyond blades, the tower; load-bearing supports; sensors for the machine and the environment; mechanical drive components, such as bearings and lubricants; and electrical drivetrain components, such as generators, as well as semiconductors used in the inverters, power-control, and grid support functions, would benefit from further innovation.

Third grand challenge: Systems science for integration of wind power plants into the future electricity grid

The global electricity system operates on several times scales, supplying all of the demand for both bulk energy and instantaneous power. Time scales vary as a function of the need for robust grid stability and reliability, operation, and planning and extend from the subsecond to decades (Fig. 5). Within each of these broader time scales, power plants must provide many functions for the grid, including protection against lightning, short circuits, and surges; robust operation under perturbation by transients, resonance, and voltage instabilities; energy demand matching within minutes to hours; and long-term predictable and controllable supply of capacity (*10*, *11*). In addition, electricity generated by large, rotating machines, such as those now found in thermal and hydroelectric plants, creates an energy transmission grid with attributes (e.g., frequency, voltage, and phase) that are defined by the physical rotation and inertia of those generators.

As physical inertia from traditional power plants decreases relative to overall system capacity, converter-based generation, such as wind and solar power plants, must provide more predictable and controllable power as well as services that support grid reliability, stability, and formation (*85*). Wind power plants today can support many of the needs of the current grid (*86*–*88*), but additional research is needed to address how wind plants of the future and their special attributes can be used to serve the demands of a converter-based grid (*12*). The path to realizing this future will require substantial research at the intersections of atmospheric flow modeling, individual turbine dynamics, wind plant control, and the larger electric system operation. The third grand challenge encompasses three intersecting research areas: wind power plant controls, the converter-dominated electric grid, and integrated data and modeling computational methods for system analysis and operation.

As a first step, researchers must solve challenges related to wind plants by providing sufficient control authority to serve an expanding set of functionalities. Growing experience with wind plants is revealing the complexity of managing systems with hundreds of stochastically driven individual wind turbine agents. Recent research highlights the possibility of not only maximizing energy production but also managing the flow field to increase

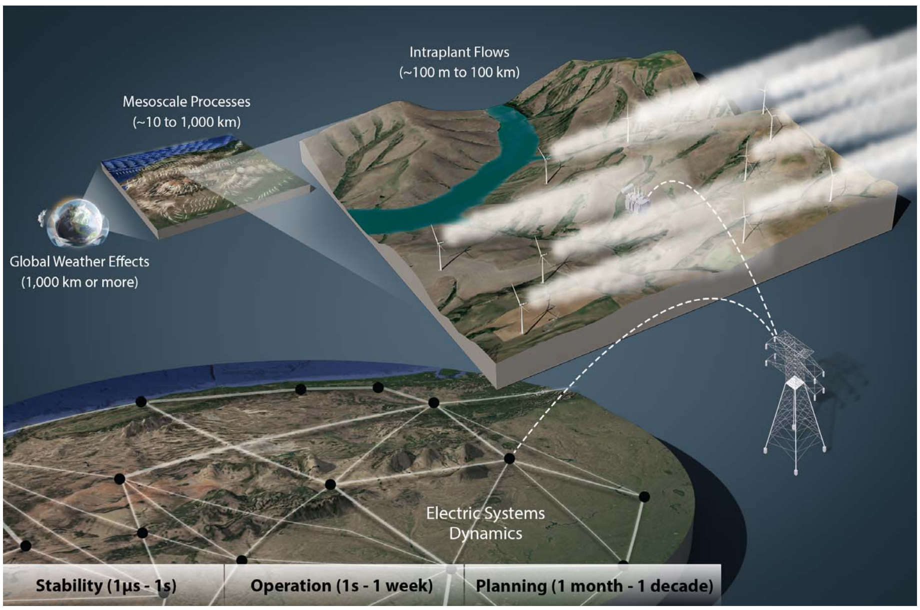


Fig. 5. Power generated by the weather-driven plant must connect to the electrical grid and support the stability, reliability, and operational needs on time scales ranging from microseconds (for managing disturbances) to decades (for long-term planning).

ILLUSTRATION: JOSH BAUER AND BESIKI KAZAISHVILI, NREL

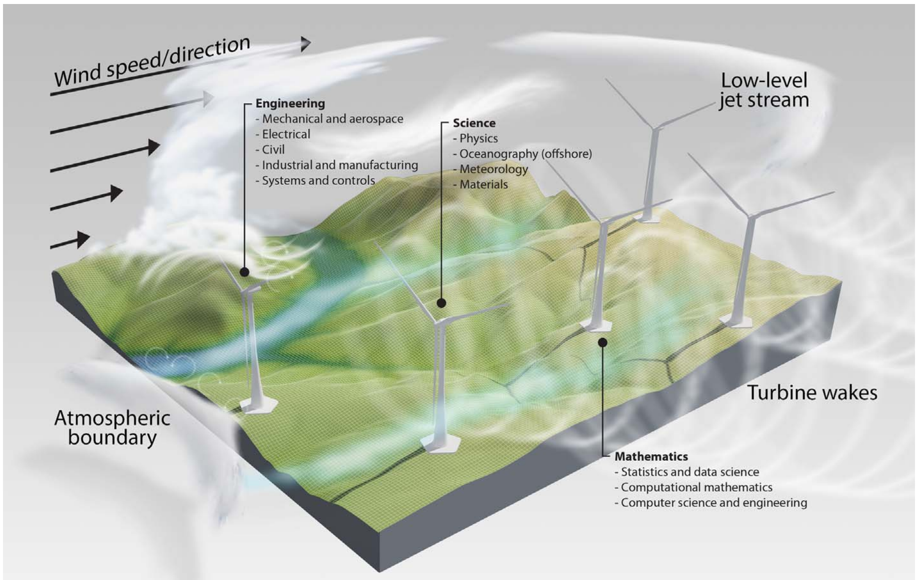


Fig. 6. A spectrum of science, engineering, and mathematics disciplines that, if integrated, can comprehensively address the grand challenges in wind energy science.

system performance (*89*–*91*). By probing the collective data available during real-time operation, new opportunities for power plant control are emerging (*92*, *93*). Greater comprehension of the wind flow and dynamics enables real-time characterization of the plant operational state and the ability to control the flow and turbine responses in the short term. Innovative controls could leverage the attributes of the machines to supply ancillary services (e.g., the rotational inertia of the blades could be tapped to ride through grid faults, or the distributed power electronics in the converters connected to the generators could be used to manage grid requirements). For example, recent work has used such integrated modeling approaches to investigate the potential for active power control from wind power plants (*86*, *94*, *95*).

The research necessary to support a future converter-dominated electric grid system extends beyond individual wind power plant controls. For example, wind turbines offer a potential source of physical inertia, but the machines (as they exist today) and solar power are typically interconnected to the grid through power electronic converters, which use software and controls to confer attributes akin to traditional power plants. Wind power plants of the future equipped with the appropriate power electronics could provide physical inertia or "synthetic inertia," with the latter enabling wind turbines to function as virtual synchronous generators (*12*, *96*, *97*). Studies that have considered up to 25% contributions of renewables to the grid need to be further refined for shares that reach beyond 50% or even 80% (*98*–*101*).

New sensors and data management techniques will also be needed to obtain and transmit real-time data on the status of the future grid, which will be governed more on information than physical inertia. Data sources will comprise a combination of measurements and simulations. Opportunities are ripe for advanced stochastic system analysis and data science that can extract meaning and direction from a combination of regional weather status and forecasts, millions of signals describing individual turbine and plant states, and real-time updates from throughout the grid. In addition, the substantial sources of uncertainty in various aspects of the system operation (from the weather-driven effects on renewable energy availability and electricity demand to availability of storage and a host of other phenomena) make this an extremely large stochastic and dynamic optimization problem that will require greater involvement by the applied mathematics and computational science communities (*98*, *101*, *102*).

A role for integrative wind energy science

These wind research grand challenges build on each other. Characterizing the wind power plant operating zone in the atmosphere will be essential to making progress in designing the next generation of even larger low-cost wind turbines, whereas understanding both dynamic control of the machines and forecasting the nature of the atmospheric inflow will enable the control of the plant necessary for grid support. Wind energy science also involves the coupling of physics across an increasingly large range of spatial and temporal scales in the atmosphere, enormous flexible aeroelastic and mechanical systems, and electrical integration with and support for a continent-sized grid system.

Although advances in individual scientific disciplines will continue to be tremendously important, recognition of the value in understanding the cross-disciplinary considerations and drivers of the technology is also paramount. In a similar way to how the aerospace discipline has driven profound achievements in materials, manufacturing, aerodynamics, structures, and controls while innovating the broader systems of aircraft and spacecraft, the emerging discipline of wind energy science seeks to leverage deep disciplinary expertise with a systems knowledge that addresses complex and multifaceted challenges.

Successful examples of integrated wind energy research are already in place at several universities and research organizations where nationally and internationally funded projects are interdisciplinary by design and aimed at tackling some of the challenges described in the preceding sections. These institutions have begun to train the next generation of scientists and engineers in departments devoted specifically to wind energy. The European Wind Energy Academy, a collaboration of more than 40 European universities with major activities in wind energy research and education, is another example of an effort to organize a scientific discipline around wind energy. Future growth of wind energy to serve global clean energy needs is expected to demand more dedicated wind energy research, cutting across the traditional disciplines. A move to embrace this shift toward studying wind energy science as its own discipline can be achieved by drawing in researchers from a range of different fields, as shown in Fig. 6.

In addition to the wide-ranging science, engineering, and mathematics needs for integrated wind energy research, methods for handling and streamlining exchange of vast quantities of information across many disciplines (both experimental and computational) will be crucial to enabling successful integrated research (*33*, *96*, *101*, *103*, *104*). Research in fields related to computational and data science will further support the wind scientific community as it seeks to integrate models and data across different scales and disciplines (*105*, *106*).

This interdisciplinary wind energy science and engineering approach offers the potential to develop solutions that not only advance the state-of-the-art in turbine subsystems but also create the integrated solutions necessary for advancing the entire system—from the turbine to the plant to the overall electrical grid. These gains are most likely to be successful when activities in a respective area are informed by a comprehensive view of the realities of the larger

context. The long-term research challenges are ripe for immediate action, and progress will depend on a generation of scientists educated deeply in their own specialty as well as in the breadth of wind energy science.

REFERENCES AND NOTES

1. UN, World Population Prospects 2017 (2017); https://population.un.org/wpp/DVD/Files/2_Indicators%20(Probabilistic%20Projections)/UN_PPP2017_Output_PopTot.xls.
2. BNEF, NEO 2018 presentation at CSIS (2018); https://about.bnef.com/blog/neo-2018-presentation-csis/.
3. International Energy Agency (IEA), "World energy outlook 2018" (Tech. Rep., IEA, 2018); www.iea.org/weo2018/.
4. BNEF, "New energy outlook 2019" (2019); https://about.bnef.com/new-energy-outlook/.
5. BP Energy Economics, "BP Energy Outlook: 2018 edition" (2018); www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/energy-outlook/bp-energy-outlook-2018.pdf.
6. DNV GL, "Energy transition outlook 2018: A global and regional forecast of the Energy transition to 2050" (2018); https://eto.dnvgl.com/2018/.
7. Global Wind Energy Council, "51.3 GW of global wind capacity installed in 2018" (2019); https://gwec.net/51-3-gw-of-global-wind-capacity-installed-in-2018/.
8. T. Mai, E. Lantz, M. Mowers, R. Wiser, "The value of wind technology innovation: Implications for the U.S. power system, wind industry, electricity consumers, and environment" (Tech. Rep. NREL/TP-6A20-70032, NREL, 2017); www.nrel.gov/docs/fy17osti/70032.pdf.
9. International Renewable Energy Agency (IRENA), "Global energy transformation: A roadmap to 2050" (IRENA, 2018); www.irena.org/-/media/Files/IRENA/Agency/Publication/2018/Apr/IRENA_Report_GET_2018.pdf.
10. M. Ahlstrom *et al*., The evolution of the market: Designing a market for high levels of variable generation. *IEEE Power Energy Mag.* **13**, 60–66 (2015). doi: 10.1109/MPE.2015.2458755
11. B. Kroposki *et al*., Achieving a 100% Renewable Grid: Operating Electric Power Systems with Extremely High Levels of Variable Renewable Energy. *IEEE Power Energy Mag.* **15**, 61–73 (2017). doi: 10.1109/MPE.2016.2637122
12. T. Ackermann *et al*., Paving the Way: A Future Without Inertia Is Closer Than You Think. *IEEE Power Energy Mag.* **15**, 61–69 (2017). doi: 10.1109/MPE.2017.2729138
13. E. Pursiheimo, H. Holttinen, T. Koljonen, Inter-sectoral effects of high renewable energy share in global energy system. *Renew. Energy* **136**, 1119–1129 (2019). doi: 10.1016/j.renene.2018.09.082
14. M. B. Milligan *et al*., Alternatives No More: Wind and Solar Power Are Mainstays of a Clean, Reliable, Affordable Grid. *IEEE Power Energy Mag.* **13**, 78–87 (2015). doi: 10.1109/MPE.2015.2462311
15. "Converter dominated" refers to a grid system largely composed of converter-based generation technologies (such as wind and solar) that convert ac to dc, which can then be fed to the larger system via dc transmission or inverted back to ac to feed a larger ac system.
16. A. Betz, *Wind-Energie und ihre Ausnutzung durch Windmuuhlen* (Univ. of Goettingen Press, 1926).
17. R. Wiser, M. Bolinger, "2017 wind technologies market report" (DOE Tech. Rep. DOE/EE-1798, 2018); www.energy.gov/eere/wind/downloads/2017-wind-technologies-market-report.
18. E. Lantz, R. Wiser, M. Hand, "IEA wind task 26: The past and future cost of wind energy" (Tech. Rep. NREL/TP-6A20-53510, NREL, 2012); www.nrel.gov/docs/fy12osti/53510.pdf.
19. Lazard, "Lazard's levelized cost of energy analysis–version 11.0" (Tech. Rep., Lazard, 2018); www.lazard.com/media/450337/lazard-levelized-cost-of-energy-version-110.pdf.
20. P. Donohoo-Vallett, "Revolution…now: The future arrives for five clean energy technologies – 2016 update" (DOE, 2016); www.energy.gov/sites/prod/files/2016/09/f33/Revolutiona%CC%82%E2%82%ACNow%202016%20Report_2.pdf.
21. P. S. Veers *et al*., Trends in the design, manufacture and evaluation of wind turbine blades. *Wind Energy* **6**, 245–259 (2003). doi: 10.1002/we.90
22. P. Jamieson, *Innovation in Wind Turbine Design* (Wiley, 2011).
23. S. Scott *et al*., Effects of aeroelastic tailoring on performance characteristics of wind turbine systems. *Renew. Energy* **114**, 887–903 (2017). doi: 10.1016/j.renene.2017.06.048
24. P. Bortolotti, C. L. Bottasso, A. Croce, L. Sartori, Integration of Multiple Passive Load Mitigation Technologies by Automated Design Optimization - The Case Study of a Medium-Size Onshore Wind Turbine. *Wind Energy* **22**, 65–79 (2019). doi: 10.1002/we.2270
25. G. van Kuik, J. Peinke, Eds. *Long-Term Research Challenges in Wind Energy – A Research Agenda by the European Academy of Wind Energy* (vol. 6 of Research Topics in Wind Energy Series, Springer, 2016).
26. K. Dykes *et al*., "Results of IEA Wind TCP workshop on a grand vision for wind energy technology" (IEA Wind TCP Task 11, Tech. Rep. NREL/TP-5000-72437, NREL, 2019); www.nrel.gov/docs/fy19osti/72437.pdf.
27. Many technology innovation pathways, including concepts such as airborne wind turbines, were discussed and documented in the International Energy Agency Grand Wind Workshop report. Progress on the grand challenges is essential to enabling such technology configurations. However, the focus of this article is on major breakthroughs, even with power plants comprising standard horizontal-axis wind turbines.
28. J. Mann, The spatial structure of neutral atmospheric surface-layer turbulence. *J. Fluid Mech.* **273**, 141–168 (1994). doi: 10.1017/S0022112094001886
29. J. C. Kaimal, J. C. Wyngaard, Y. Izumi, O. R. Coté, Spectral characteristics of surface-layer turbulence. *Q. J. R. Meteorol. Soc.* **98**, 563–589 (1972). doi: 10.1002/qj.49709841707
30. P. S. Veers, "Three-dimensional wind simulation" (Report no. SAND-88-0152C, CONF-890102-9, Sandia National Laboratories, 1988); https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/1988/880152.pdf.
31. J. C. Wyngaard, Toward Numerical Modeling in the "Terra Incognita". *J. Atmos. Sci.* **61**, 1816–1826 (2004). doi: 10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2
32. X. G. Larsén, E. L. Petersen, S. E. Larsen, Variation of boundary-layer wind spectra with height. *Q. J. R. Meteorol. Soc.* **144**, 2054–2066 (2018). doi: 10.1002/qj.3301
33. J. S. Sanz Rodrigo *et al*., Mesoscale to microscale wind farm modeling and evaluation. *WIREs Energy Environ.* **6**, e214 (2017). doi: 10.1002/wene.214
34. J. Mann *et al*., Complex terrain experiments in the New European Wind Atlas. *Philos. Trans. R. Soc. A* **375**, 20160101 (2017). doi: 10.1098/rsta.2016.0101; pmid: 28265025
35. D. A. Rajewski *et al*., Crop Wind Energy Experiment (CWEX): Observations of Surface-Layer, Boundary Layer, and Mesoscale Interactions with a Wind Farm. *Bull. Am. Meteorol. Soc.* **94**, 655–672 (2013). doi: 10.1175/BAMS-D-11-00240.1
36. X. Han, D. Liu, C. Xu, W. Z. Shen, Atmospheric stability and topography effects on wind turbine performance and wake properties in complex terrain. *Renew. Energy* **126**, 640–651 (2018). doi: 10.1016/j.renene.2018.03.048
37. J. A. Lee *et al*., Improving Wind Predictions in the Marine Atmospheric Boundary Layer through Parameter Estimation in a Single-Column Model. *Mon. Weather Rev.* **145**, 5–24 (2017). doi: 10.1175/MWR-D-16-0063.1
38. M. Calaf, C. Meneveau, J. Meyers, Large eddy simulations of fully developed wind-turbine array boundary layers. *Phys. Fluids* **22**, 015110 (2010). doi: 10.1063/1.3291077
39. N. G. Nygaard, A. C. Newcombe, Wake behind an offshore wind farm observed with dual-Doppler radars. *J. Phys. Conf. Ser.* **1037**, 072008 (2018); .doi: 10.1088/1742-6596/1037/7/072008
40. R. J. A. M. Stevens, C. Meneveau, Flow Structure and Turbulence in Wind Farms. *Annu. Rev. Fluid Mech.* **49**, 311–339 (2017). doi: 10.1146/annurev-fluid-010816-060206
41. D. W. Keith *et al*., The influence of large-scale wind power on global climate. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16115–16120 (2004). doi: 10.1073/pnas.0406930101; pmid: 15536131
42. A. C. Fitch *et al*., Local and Mesoscale Impacts of Wind Farms as Parameterized in a Mesoscale NWP Model. *Mon. Weather Rev.* **140**, 3017–3038 (2012). doi: 10.1175/MWR-D-11-00352.1
43. P. J. H. Volker, J. Badger, A. N. Hahmann, S. Ott, The Explicit Wake Parametrisation V1.0: A wind farm parametrisation in the mesoscale model WRF. *Geosci. Model Dev.* **8**, 3715–3731 (2015). doi: 10.5194/gmd-8-3715-2015
44. A. Armstrong *et al*., Ground-level climate at a peatland wind farm in Scotland is affected by wind turbine operation. *Environ. Res. Lett.* **11**, 044024 (2016). doi: 10.1088/1748-9326/11/4/044024
45. J. C. Y. Lee, J. K. Lundquist, Evaluation of the wind farm parameterization in the Weather Research and Forecasting model (version 3.8.1) with meteorological and turbine power data. *Geosci. Model Dev.* **10**, 4229–4244 (2017). doi: 10.5194/gmd-10-4229-2017
46. A. Platis *et al*., First in situ evidence of wakes in the far field behind offshore wind farms. *Sci. Rep.* **8**, 2163 (2018). doi: 10.1038/s41598-018-20389-y; pmid: 29391440
47. S. K. Siedersleben *et al*., Micrometeorological Impacts of Offshore Wind Farms as seen in Observations and Simulations. *Environ. Res. Lett.* **13**, 124012 (2018). doi: 10.1088/1748-9326/aaea0b
48. M. Z. Jacobson, C. L. Archer, Saturation wind power potential and its implications for wind energy. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15679–15684 (2012). doi: 10.1073/pnas.1208993109; pmid: 23019353
49. A. S. Adams, D. W. Keith, Are global wind power resource estimates overstated? *Environ. Res. Lett.* **8**, 015021 (2013). doi: 10.1088/1748-9326/8/1/015021
50. A. C. Fitch, J. K. Lundquist, J. B. Olson, Mesoscale Influences of Wind Farms throughout a Diurnal Cycle. *Mon. Weather Rev.* **141**, 2173–2198 (2013). doi: 10.1175/MWR-D-12-00185.1
51. J. K. Lundquist, K. K. DuVivier, D. Kaffine, J. M. Tomaszewski, Costs and consequences of wind turbine wake effects arising from uncoordinated wind energy development. *Nat. Energy* **4**, 26–34 (2019). doi: 10.1038/s41560-018-0281-2
52. K. B. Karnauskas, J. K. Lundquist, L. Zhang, Southward shift of the global wind energy resource under high carbon dioxide emissions. *Nat. Geosci.* **11**, 38–43 (2018). doi: 10.1038/s41561-017-0029-9
53. C. L. Archer *et al*., Meteorology for coastal/offshore wind energy in the United States: Recommendations and research needs for the next 10 years. *Bull. Am. Meteorol. Soc.* **95**, 515–519 (2014). doi: 10.1175/BAMS-D-13-00108.1
54. C. Meneveau, Big wind power: Seven questions for turbulence research. *J. Turbul.* **20**, 2–20 (2019). doi: 10.1080/14685248.2019.1584664
55. M. L. Aitken, R. M. Banta, Y. L. Pichugina, J. K. Lundquist, Quantifying wind turbine wake characteristics from scanning remote sensor data. *J. Oceanic Atmos. Technol.* **31**, 765–787 (2014). doi: 10.1175/JTECH-D-13-00104.1
56. A. Peña, C. B. Hasager, M. Badger, R. J. Barthelmie, F. Bingöl, J.-P. Cariou, S. Emeis, S. T. Frandsen, M. Harris, I. Karagali, S. E. Larsen, J. Mann, T. Mikkelsen, M. Pitter, S. C. Pryor, A. Sathe, D. Schlipf, C. Slinger, R. Wagner, "Remote sensing for wind energy" (DTU Wind Energy-E-Report-0084, Denmark Technical University, 2015); https://orbit.dtu.dk/files/111814239/DTU_Wind_Energy_Report_E_0084.pdf.
57. A. Clifton *et al*., IEA Wind Task 32: Wind Lidar Identifying and Mitigating Barriers to the Adoption of Wind Lidar. *Remote Sens.* **10**, 406 (2018). doi: 10.3390/rs10030406
58. B. D. Hirth, J. L. Schroeder, W. S. Gunter, J. G. Guynes, Measuring a Utility-Scale Turbine Wake Using the TTUKa Mobile Research Radars. *J. Atmos. Ocean. Technol.* **29**, 765–771 (2012). doi: 10.1175/JTECH-D-12-00039.1
59. R. Menke, N. Vasiljević, K. S. Hansen, A. N. Hahmann, J. Mann, Does the wind turbine wake follow the topography? A multi-lidar study in complex terrain. *Wind Energy Sci.* **3**, 681–691 (2018). doi: 10.5194/wes-3-681-2018
60. N. Wildmann, N. Vasiljevic, T. Gerz, Wind turbine wake measurements with automatically adjusting scanning trajectories in a multi-Doppler lidar setup. *Atmos. Meas. Tech.* **11**, 3801–3814 (2018). doi: 10.5194/amt-11-3801-2018
61. C. B. Hasager *et al*., Remote Sensing Observation Used in Offshore Wind Energy. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **1**, 67–79 (2008). doi: 10.1109/JSTARS.2008.2002218
62. A. M. Sempreviva, R. J. Barthelmie, S. C. Pryor, Review of Methodologies for Offshore Wind Resource Assessment in European Seas. *Surv. Geophys.* **29**, 471–497 (2008). doi: 10.1007/s10712-008-9050-2
63. S. E. Haupt, R. Kotamarthi, Y. Feng, J. D. Mirocha, E. Koo, R. Linn, B. Kosovic, B. Brown, A. Anderson, M. J. Churchfield, C. Draxl, E. Quon, W. Shaw, L. Berg, R. Rai, B. L. Ennis, "Second year report of the atmosphere to electrons mesoscale to microscale coupling project: Nonstationary modeling techniques and assessment" (PNNL-26267, Pacific Northwest National Laboratory, 2017); www.pnnl.gov/main/publications/external/technical_reports/PNNL-26267.pdf.
64. S. Lee, M. Churchfield, P. Moriarty, J. Jonkman, J. Michalakes, "Atmospheric and wake turbulence impacts on wind turbine fatigue loadings: Preprint" (Conference Paper NREL/CP-5000-53567, NREL, 2011); www.nrel.gov/docs/fy12osti/53567.pdf.
65. H. A. Madsen, N. N. Sørensen, C. Bak, N. Troldborg, G. Pirrung, Measured aerodynamic forces on a full scale 2MW turbine in comparison with EllipSys3D and HAWC2 simulations. *J. Phys. Conf. Ser.* **1037**, 022011 (2018). doi: 10.1088/1742-6596/1037/2/022011

66. J. G. Schepers *et al*., Final results from the EU project AVATAR: Aerodynamic modelling of 10 MW wind turbines. *J. Phys. Conf. Ser.* **1037**, 022013 (2018). doi: 10.1088/1742-6596/1037/2/022013
67. M. A. Sprague, S. Boldyrev, P. Fischer, R. Grout, W. I. Gustafson Jr., R. Moser, "Turbulent flow simulation at the exascale: Opportunities and challenges workshop," 4 to 5 August 2015 (NREL/TP-2C00-67648, DOE, 2017); www.nrel.gov/docs/fy17osti/67648.pdf.
68. J. C. Heinz, N. N. Sørensen, F. Zahle, Fluid-structure interaction computations for geometrically resolved rotor simulations using CFD. *Wind Energy* **19**, 2205–2221 (2016). doi: 10.1002/we.1976
69. A. R. Stäblein, M. H. Hansen, G. Pirrung, Fundamental aeroelastic properties of a bend–twist coupled blade section. *J. Fluids Structures* **68**, 72–89 (2017). doi: 10.1016/j.jfluidstructs.2016.10.010
70. A. R. Stäblein, M. H. Hansen, D. R. Verelst, Modal Properties and Stability of Bend-Twist Coupled Wind Turbine Blades. *Wind Energy Sci.* **2**, 343–360 (2017). doi: 10.5194/wes-2-343-2017
71. R. Riva, S. Cacciola, C. L. Bottasso, Periodic Stability Analysis of Wind Turbines Operating in Turbulent Wind Conditions. *Wind Energy Sci.* **1**, 177–203 (2016). doi: 10.5194/wes-1-177-2016
72. A. Morató, S. Sriramula, N. Krishnan, J. Nichols, Ultimate loads and response analysis of a monopile supported offshore wind turbine using fully coupled simulation. *Renew. Energy* **101**, 126–143 (2017). doi: 10.1016/j.renene.2016.08.056
73. L. Suja-Thauvin, J. R. Krokstad, E. E. Bachynski, Critical assessment of non-linear hydrodynamic load models for a fully flexible monopile offshore wind turbine. *Ocean Eng.* **164**, 87–104 (2018). doi: 10.1016/j.oceaneng.2018.06.027
74. J. Jonkman, Dynamics of Offshore Floating Turbines – Model Development and Verification. *Wind Energy* **12**, 459–492 (2009). doi: 10.1002/we.347
75. J. Jose, S. J. Choi, K. E. Giljarhus, O. T. Gudmestad, A comparison of numerical simulations of breaking wave forces on a monopile structure using two different numerical models based on finite difference and finite volume methods. *Ocean Eng.* **137**, 78–88 (2017). doi: 10.1016/j.oceaneng.2017.03.045
76. A. Natarajan, Influence of second-order random wave kinematics on the design loads of offshore wind turbine support structures. *Renew. Energy* **68**, 829–841 (2014). doi: 10.1016/j.renene.2014.02.052
77. B. T. Paulsen, H. Bredmose, H. B. Bingham, N. G. Jacobsen, Forcing of a bottom-mounted circular cylinder by steep regular water waves at finite depth. *J. Fluid Mech.* **755**, 1–34 (2014). doi: 10.1017/jfm.2014.386
78. R. Worsnop, J. K. Lundquist, G. H. Bryan, R. Damiani, W. Musial, Gusts and Shear Within Hurricane Eyewalls Can Exceed Offshore Wind-Turbine Design Standards. *Geophys. Res. Lett.* **44**, 6413–6420 (2017). doi: 10.1002/2017GL073537
79. E. Kim, L. Manuel, M. Curcic, S. S. Chen, C. Phillips, P. Veers, "On the use of coupled wind, wave, and current fields in the simulation of loads on bottom-supported offshore wind turbines during hurricanes" (DOE/GO-102015-4798 2016, Tech. Rep. NREL/TP -5000-65283, NREL, 2016); www.nrel.gov/docs/fy16osti/65283.pdf.
80. T. Han, G. McCann, T. A. Mücke, K. Freudenreich, How can a wind turbine survive a tropical cyclone? *Renew. Energy* **70**, 3–10 (2014). doi: 10.1016/j.renene.2014.02.014
81. J. N. Sørensen, W. Z. Shen, X. Munduate, Analysis of wake states by a full-field actuator disc model. *Wind Energy* **1**, 73–88 (1998). doi: 10.1002/(SICI)1099-1824(199812)1:2<73::AID-WE12>3.0.CO;2-L
82. C. Lienard, R. Boisard, C. Daudin, "Aerodynamic behavior of a floating offshore wind turbine," AIAA SciTech 2019 Forum, San Diego, CA, 7 to 11 January 2019; doi: 10.2514/6.2019-1575
83. B. Koo, A. J. Goupee, R. W. Kimball, K. F. Lambrakos, Model Tests for a Floating Wind Turbine on Three Different Floaters. *J. Offshore Mech. Arctic Eng.* **136**, 020907 (2014). doi: 10.1115/1.4024711
84. D. S. Cousins, Y. Suzuki, R. E. Murray, J. R. Samaniuk, A. P. Stebner, Recycling glass fiber thermoplastic composites from wind turbine blades. *J. Clean. Prod.* **209**, 1252–1263 (2019). doi: 10.1016/j.jclepro.2018.10.286
85. Grid formation involves supporting the fundamental structure of an electric grid system. This includes serving as a reliable voltage source for ac or dc systems and providing frequency signals for ac systems.
86. C. Shapiro, P. Bauweraerts, J. Meyers, C. Meneveau, D. F. Gayme, Model-based receding horizon control of wind farms for secondary frequency regulation. *Wind Energy* **20**, 1261–1275 (2017). doi: 10.1002/we.2093
87. V. Gevorgian, Y. Zhang, E. Ela, Investigating the Impacts of Wind Generation Participation in Interconnection Frequency Response. *IEEE Trans. Sustainable Energy* **6**, 1004–1012 (2015). doi: 10.1109/TSTE.2014.2343836
88. J. W. van Wingerden, L. Y. Pao, J. Aho, P. Fleming, "Active Power Control of Waked Wind Farms." *Proc. IFAC World Congress*, Toulouse, France, (2017), pp. 4570–4577. doi: 10.1016/j.ifacol.2017.08.378
89. P. M. O. Gebraad *et al*., Wind plant power optimization through yaw control using a parametric model for wake effects – a CFD simulation study. *Wind Energy* **19**, 95–114 (2014). doi: 10.1002/we.1822
90. P. Fleming *et al*., Full-Scale Field Test of Wake Steering. *J. Phys. Conf. Ser.* **854**, 012013 (2017). doi: 10.1088/1742-6596/854/1/012013
91. P. Fleming *et al*., A simulation study demonstrating the importance oflarge-scale trailing vortices in wake steering. *Wind Energy Sci.* **3**, 243–255 (2018). doi: 10.5194/wes-3-243-2018
92. J. Annoni *et al*., A Wind direction estimation using SCADA data with consensus-based optimization. *Wind Energy Sci.* **4**, 355–368 (2019). doi: 10.5194/wes-2018-60
93. S. Raach, S. Boersma, B. Doekemeijer, J.-W. van Wingerden, P. W. Cheng, Lidar-based closed-loop wake redirection in high-fidelity simulation. *J. Phys. Conf. Ser.* **1037**, 032016 (2018). doi: 10.1088/1742-6596/1037/3/032016
94. P. A. Fleming, J. Aho, P. Gebraad, L. Pao, Y. Zhang, "Computational fluid dynamics simulation study of active power control in wind plants" in *Proc. American Control Conf.* (IEEE, 2016), pp. 1413–1420; .doi: 10.1109/ACC.2016.7525115
95. M. Vali, V. Petrović, G. Steinfeld, L. Y. Pao, M. Kühn, An active power control approach for wake-induced load alleviation in a fully developed wind farm boundary layer. *Wind Energy Sci.* **4**, 139–161 (2019). doi: 10.5194/wes-4-139-2019
96. D. Pattabiraman, R. H. Lasseter, T. M. Jahns, "Comparison of grid following and grid forming control for a high inverter penetration power system," in *2018 IEEE Power Energy Society General Meeting* (IEEE, 2018); .doi: 10.1109/PESGM.2018.8586162
97. P. Fairley, "Can synthetic inertia from wind power stabilize grids?" *IEEE Spectrum* (2016); https://spectrum.ieee.org/energywise/energy/renewables/can-synthetic-inertia-stabilize-power-grids.
98. Q. P. Zheng, J. Wang, A. L. Liu, Stochastic Optimization for Unit Commitment—A Review. *IEEE Trans. Power Syst.* **30**, 1913–1924 (2015). doi: 10.1109/TPWRS.2014.2355204
99. D. Molzahn *et al*., A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems. *IEEE Trans. Smart Grid* **8**, 2941–2962 (2017). doi: 10.1109/TSG.2017.2720471
100. H. Holttinen *et al*., "Design and operation of power systems with large amounts of wind power" (Final summary report, IEA Wind Task 25, IEA, 2016); www.vtt.fi/inf/pdf/technology/2016/T268.pdf.
101. N. Helistö, J. Kiviluoma, H. Holttinen, J. D. Lara, B.-M. Hodge, Including operational aspects in the planning of power systems with large amounts of variable generation: A review of modelling approaches. *WIREs Energy Environ.* **8**, e341 (2019). doi: 10.1002/wene.341
102. L. Göransson, F. A. Johnsson, A comparison of variation management strategies for wind power integration in different electricity system contexts. *Wind Energy* **21**, 837–854 (2018). doi: 10.1002/we.2198
103. M. D. Wilkinson *et al*., The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016). doi: 10.1038/sdata.2016.18; pmid: 26978244
104. A. M. Sempreviva *et al*., Taxonomy and metadata for wind energy research & development. Version 1, Zenodo (2017). doi: 10.5281/ZENODO.1199489
105. European Technology & Innovation Platform on Wind Energy (ETIP Wind) Steering Committee, "When wind goes digital" (ETIP Wind, 2014); https://etipwind.eu/news/wind-goes-digital/.
106. C. E. Concolato, L. M. Chen, Data science: A new paradigm in the age of big-data science and analytics. *New Math. Nat. Computation* **13**, 119–143 (2017). doi: 10.1142/S1793005717400038

ACKNOWLEDGMENTS

We thank the IEA Wind leadership and membership for supporting and participating in the "Grand Vision for Wind Energy" workshop that inspired the effort leading to this article. In addition, we thank the authors of the compendium that stemmed from the workshop, which was an important source of comprehensive background information on the state of the art in wind energy science research and trends in wind energy innovation. We also thank senior management at NREL and the U.S. Department of Energy (DOE) for supporting the article development and review. Editorial assistance at NREL was provided by S. Anstedt. **Funding:** This work was authored (in part) by NREL, operated by Alliance for Sustainable Energy, LLC, for the DOE under contract DE-AC36-08GO28308. Funding for NREL is provided by the DOE Office of Energy Efficiency and Renewable Energy. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. government. The U.S. government retains (and the publisher, by accepting the article for publication, acknowledges that the U.S. government retains) a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. government purposes. L.P. acknowledges support from a Palmer Endowed Chair Professorship and a fellowship from the Hanse-Wissenschaftskolleg Institute for Advanced Study (Delmenhorst, Germany). The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect the views of NOAA or the Department of Commerce. **Competing interests:** C.M. serves on the Scientific Advisory Board of the Max Planck Institute for Self-Organization (Goettingen, Germany), where some of the scientists are involved in wind energy research; he also serves as a paid expert witness in a legal case involving a wind farm in the United States. J.C.S. is the director of Nexgen Energy, LLC, a consulting company. No other authors have competing interests in commercial activities related to wind energy that are not already explicitly clear from their affiliations.

10.1126/science.aau2027

REVIEW SUMMARY

MICROBIOTA

Vulnerability of the industrialized microbiota

Justin L. Sonnenburg* and Erica D. Sonnenburg*

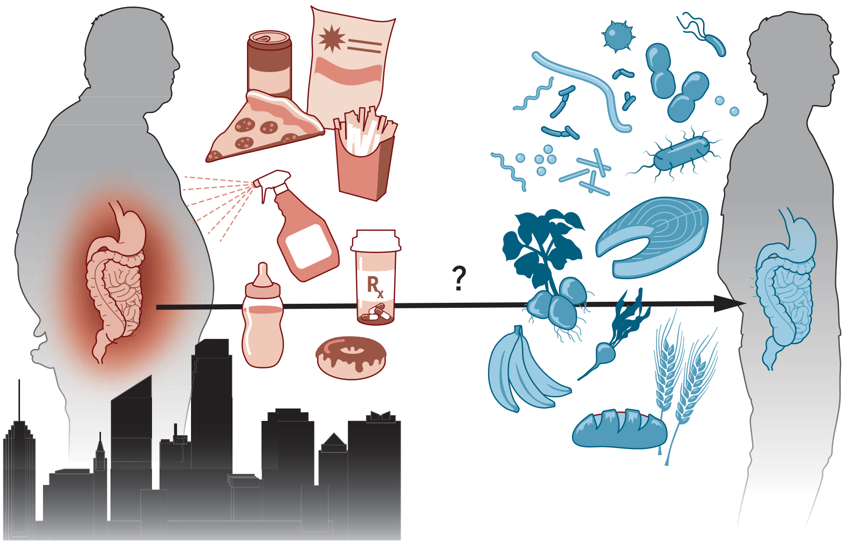
BACKGROUND: The collection of trillions of microbes inhabiting the human gut, called the microbiome or microbiota, has captivated the biomedical research community for the past decade. Intimate connections exist between the microbiota and the immune system, central nervous system, and metabolism. The growing realization of the fundamental role that the microbiota plays in human health has been accompanied by the challenge of trying to understand which features define a healthy gut community and how these may differ depending upon context. Such insight will lead to new routes of disease treatment and prevention and may illuminate how lifestyle-driven changes to the microbiota can impact health across populations. Individuals living traditional lifestyles around the world share a strikingly similar microbiota composition that is distinct from that found in industrialized populations. Indeed, lineages of gut microbes have cospeciated with humans over millions of years, passing through hundreds of thousands of generations, and lend credence to the possibility that our microbial residents have shaped our biology throughout evolution. Relative to the "traditional" microbiota, the "industrial" microbiota appears to have lower microbial diversity, with major shifts in membership and functions. Individuals immigrating from nonindustrialized to industrialized settings or living at different intermediate states between foraging and industrialization have microbiota composition alterations that correspond to time and severity of lifestyle change. Industrial advances including antibiotics, processed food diets, and a highly sanitized environment have been shown to influence microbiota composition and transmission and were developed and widely implemented in the absence of understanding their effects on the microbiota.

ADVANCES: Here, we argue that the microbiota harbored by individuals living in the industrialized world is of a configuration never before experienced by human populations. This "new," industrial microbiota has been shaped by recent progress in medicine, food, and sanitation. As technology and medicine have limited our exposure to pathogenic microbes, enabled feeding large populations inexpensively, and otherwise reduced acute medical incidents, many of these advances have been implemented in the absence of understanding the collateral damage inflicted on our resident microbes or the importance of these microbes in our health. More connections are being drawn between the composition and function of the gut microbiota and alteration in the immune status of the host. These relationships connect the industrial microbiota to the litany of chronic diseases that are driven by inflammation. Notably, these diseases spread along with the lifestyle factors that are known to alter the microbiota. While researchers have been uncovering the basic tenets of how the microbiota influences human health, there has been a growing realization that as the industrial lifestyle spreads globally, changes to the human microbiota may be central to the coincident spread of non-communicable, chronic diseases and may not be easily reversed.

OUTLOOK: We suggest that viewing microbiota biodiversity with an emphasis on sustainability and conservation may be an important approach to safeguarding human health. Understanding the services provided by the microbiota to humans, analogous to how ecosystem services are used to place value on aspects of macroecosystems, could aid in assessing the cost versus benefit of specific microbiota dysfunctions that are induced by different aspects of lifestyle. A key hurdle is to establish the impact of industrialization-induced changes to the microbiota on human health. The severity of this impact might depend on the specifics of numerous factors, including health status, diet, human genotype, and lifestyle. Isolating and archiving bacterial strains that are sensitive to industrialization may be required to enable detailed study of these organisms and to preserve ecosystem services that are unique to those strains and potentially beneficial to human health. Determining a path forward for sustainable medical practices, diet, and sanitation that is mindful of the importance and fragility of the microbiota is needed if we are to maintain a sustainable relationship with our internal microbial world. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaw9255



Industrialization affects the human gut microbiota. Aspects of lifestyle, including those associated with industrialization, such as processed foods, infant formula, modern medicines, and sanitation, can change the gut microbiota. Major questions include whether microbiota changes associated with industrialization are important for human health, if they are reversible, and what steps should be taken to prevent further change while information is acquired to enable an informed cost-versus-benefit analysis. It is possible that a diet rich in whole foods and low in processed foods, along with increased exposure to nonpathogenic microbes, may be beneficial to industrial populations.

ILLUSTRATION: KELLIE HOLOSKI/*SCIENCE*

The list of author affiliations is available in the full article online.
*Corresponding author. Email: jsonnenburg@stanford.edu (J.L.S.); erica.sonnenburg@stanford.edu (E.D.S.)
Cite this article as J. L. Sonnenburg and E. D. Sonnenburg, *Science* **366**, eaaw9255 (2019). DOI: 10.1126/science.aaw9255

REVIEW

MICROBIOTA

Vulnerability of the industrialized microbiota

Justin L. Sonnenburg[1,2]* and Erica D. Sonnenburg[1]*

The human body is an ecosystem that is home to a complex array of microbes known as the microbiome or microbiota. This ecosystem plays an important role in human health, but as a result of recent lifestyle changes occurring around the planet, whole populations are seeing a major shift in their gut microbiota. Measures meant to kill or limit exposure to pathogenic microbes, such as antibiotics and sanitation, combined with other factors such as processed food, have had unintended consequences for the human microbial ecosystem, including changes that may be difficult to reverse. Microbiota alteration and the accompanying loss of certain functional attributes might result in the microbial communities of people living in industrialized societies being suboptimal for human health. As macroecologists, conservationists, and climate scientists race to document, understand, predict, and delay global changes in our wider environment, microbiota scientists may benefit by using analogous approaches to study and protect our intimate microbial ecosystems.

Ecosystems change. Seasonal or periodic fluctuations may occur over short time scales, trajectories of lasting change can occur over time, and sudden perturbations can result in instability or new stable states. Ecosystems can also reach tipping points at which biodiversity crashes, invasive and opportunistic species take over, and the services expected of the original ecosystem are lost, which may result in further damage and/or extinctions. Each human is an ecosystem composed of thousands of species and trillions of members, the host body of *Homo sapiens* being just one of those species. Most of these community members are microorganisms that reside in the gut, which is the focus of this article. Sequencing of the microbiota shows that human microbiomes are composed of a stunning array of species and functional diversity. An intricate set of interactions, just now being mapped, connects microbial species within a microbiota to one another and to human biology and is beginning to show how profoundly these microbes influence our health.

The first steps in human microbiota assembly occur upon birth, with microbes vying to colonize environment-exposed surfaces in and on the body. This process is influenced by many factors, including mode of birth, nutrition, environment, infection, and antibiotic exposure (*1*, *2*). Specific taxa of microbes have codiversified with *Homo sapiens*, consistent with vertical transmission over hundreds of thousands of generations (*3*). The millions of years of association have provided ample opportunities for our biology and theirs to coevolve (*4*).

[1]Department of Microbiology and Immunology and Center for Human Microbiome Studies, Stanford University, Stanford, CA, USA. [2]Chan Zuckerberg Biohub, San Francisco, CA, USA.
*Corresponding author. Email: jsonnenburg@stanford.edu (J.L.S.); erica.sonnenburg@stanford.edu (E.D.S.)

Intimate connections between the microbiota and the human immune system, nervous system, and metabolism have been revealed over the past decade (*5*–*9*). The specific microbes that first colonize the infant gut and the ensuing succession of the community can irreversibly influence mucosal and systemic immune development (*10*). Orchestrating the assembly of a health-promoting gut microbiota or manipulating a mature community to alter human physiology has vast therapeutic potential, which has captured the attention of the biomedical community. Beyond the importance of the microbiota to human health, recent research has also demonstrated its vulnerability. This ecosystem is susceptible to change by selective forces (*11*, *12*). For example, a single course of one type of antibiotic can decimate and reshape the gut microbiota (*13*). Exciting research is racing to identify disease treatments using microbiome manipulation, but less focus has been placed on how to protect the microbiota from damage that may be deleterious to human health (*14*).

The germ theory of disease, formalized in the 1860s by Louis Pasteur, portrayed microbes as an enemy to be controlled and eradicated. The subsequent "war" on microbes deploying hand washing, sterile surgical techniques, and antibiotics has saved countless lives. In 1900, pneumonia, tuberculosis, and infectious enteritis were the three leading causes of mortality in the United States, accounting for almost one-third of all deaths (*15*). By the end of the millennium, these infectious disease killers were replaced by chronic diseases, including heart disease, cancer, and stroke, which offered evidence of our ability to effectively manage "germs." However, the inverse relationship of infectious and chronic disease rates may share a similar underlying cause. Consistent with tenets of the hygiene hypothesis, limited exposure to microbes may result in defects in immune function and/or regulation, leading to an increasing burden of allergic and autoimmune diseases. In light of our new knowledge about the role of the microbiota in health, the war on microbes likely needs to be reconsidered in less combative terms. The profound success of germ-killing techniques and drugs developed in the past century that have minimal acute side effects has led to overuse. The rise of superbugs that are resistant to antibiotics and chemical bactericides reveals that there is a cost to our war on microbes (*16*). However, the longer-term and less obvious costs to human health of disrupting the microbiota may come from chronic metabolic and immune diseases. Although intimate, the communities that live in our guts are hard to study, and at present we do not fully understand the health impact of the differences in the microbiota observed between human populations.

Microbiota composition, diversity, and gene content in industrialized peoples vary substantially from that of more traditional rural populations and likely from that of our ancient ancestors, indicating that aspects of our lifestyle are changing our resident microbes (*4*, *17*–*20*). Antibiotics are not the only potential contributor to this effect. Other recent changes in practice, including Caesarean section (C-section) delivery, infant formula, and consumption of industrially produced foods, have all been shown to influence the gut microbiota of humans (*21*–*23*). Although these technological and medical advances have had undeniable benefits (especially for emergency health care), their implementation and widespread use have occurred without an understanding of their impact on our resident microbial communities. At one extreme, microbiota shifts coincident with industrialization may have no impact (or even a beneficial impact, for example, by removing or reducing microbes with pathogenic potential) on human health and longevity. At the other extreme, the microbiota alterations observed in industrialized populations may be a major contributor to the misregulation of the human immune system that drives chronic inflammation (*4*, *24*). Noncommunicable diseases (NCDs), such as stroke, heart disease, some cancers, chronic kidney disease, diabetes, and dementias, all of which are fueled by chronic inflammation, are associated with the worldwide expansion of industrialized lifestyles and are predicted to create a global health crisis in the coming century (*25*, *26*).

In many ways, the rapid changes experienced by the microbiota of urban humans are analogous to those observed in macroecosystems throughout the world (*27*). Over time and with tremendous efforts to generate and analyze data, a global scientific consensus has emerged that human-induced climate change

will have a devastating impact on Earth's species and ecosystems if not curtailed and reversed (28, 29). Likewise, as we become increasingly cognizant of the importance of the microbiota in dictating the duration and extent of our health, it is vital that we reframe our relationship with microbes and use strategies similar to the sustainability and biodiversity conservation efforts under way around the globe. What steps should we take now to protect resident microbes, given the current data and range of possible outcomes?

The gut: A rapidly changing ecosystem

That the gut ecosystem would change in response to marked lifestyle alterations is not surprising. What is notable is that the microbiota of traditional populations share taxa that have been lost or reduced in individuals living in the industrialized world, which we have termed "VANISH" (volatile and/or associated negatively with industrialized societies of humans) taxa (Fig. 1A) (30). A study comparing the industrialized microbiota with that of three Nepalese populations living on a gradient from foraging to farming showed the shift in microbiota composition that takes place as populations depart from a foraging lifestyle (31). Intermediate states of lifestyle change toward urbanization are accompanied by less extreme but evident changes in the microbiota (Fig. 1, B and C).

Similarly, a longitudinal study of individuals immigrating from a Thai refugee camp to the United States showed a loss of VANISH taxa within months of immigrating (32). The longer the immigrants lived in the United States, the more profound the changes. In addition to changes in microbial membership, functional differences in the microbiota correspond to lifestyle. Traditional populations such as the Hadza, a hunter-gatherer group living in Tanzania, like the immigrants from Southeast Asia, harbor microbiota with a larger and more diverse collection of carbohydrate active enzymes (CAZymes) than their industrial counterparts. CAZymes digest complex plant polysaccharides, characteristic of traditional dietary fiber intake (32, 33). By comparison, the microbiota of U.S. residents are enriched in CAZymes that degrade host mucus, which serves as a backup food source for gut microbes when dietary fiber is limited, a hallmark of the industrialized diet (33, 34). The selection of mucus-utilizing bacteria in industrialized populations is evident in the enrichment of *Akkermansia muciniphila* (a mucin-loving bacterium in the phylum Verrucomicrobia) that was found in a worldwide comparison of industrialized and nonindustrialized microbiomes (Fig. 1A) (33). Whether the loss or reduction of VANISH taxa cause or contribute to the growing burden of NCDs in humans remains to be determined. However, determining the potential importance of VANISH taxa to human biology will require efforts to maintain their diversity before it is lost (35, 36).

An ecosystem vulnerable to industrialization

We must not forget how the attempted eradication of pathogenic microbes with antibiotics, increased sanitation, and medicalized birth has saved countless lives. Other features of industrialized life, such as the Western diet and infant formula, have added convenience, increased human productivity and met the food demands of a growing population. The development and widespread implementation of these technological advances occurred before there was an understanding of their effect on the microbiota and the significance of the microbiota to human health. One difficulty in understanding the effects of different aspects of industrialization on the human gut microbiota is that so many lifestyle factors covary. Below, we summarize studies that have sought to disentangle facets of the industrialized lifestyle that change the microbiota.

The development and use of antibiotics have accompanied human population growth, industrialization, and rapid technological advances. Antibiotics have become the prototypic factor associated with industrialization that negatively affects the gut microbiota. Antibiotic resistance and increased susceptibility to enteric pathogens are well-known negative effects of antibiotic use. Accumulating data also show that oral antibiotic use has long-term effects on the composition of the gut microbiota (37). Just 5 days of ciprofloxacin was shown to decimate the gut microbial community, which only recovered slowly over the ensuing weeks and months (13). Recoveries were individualized, were incomplete, and differed in their kinetics (13). Similarly, other studies have shown that antibiotics can have a long-term impact on the microbiota—perhaps we should not be surprised because most of these medicines were originally designed to have broad-spectrum effects (38).

For most of human existence, humans consumed food and water laden with microbes, some of which caused disease. But humans also routinely consumed benign bacteria, both through incidental environmental exposure (e.g., from dirt or unsanitized food or on the skin) and from fermented foods (39). The recent shift to consuming largely sterile food and water has likely also influenced the microbiota. For example, the source of drinking water was significantly associated with microbiota composition in the cross-sectional study of Nepalese individuals living on a lifestyle gradient, as well as the Hadza (31). As industrial populations removed microbes from drinking water, the burden of diseases such as cholera and other waterborne illnesses decreased. Recent studies in mice suggest that sanitization in the form of cage cleaning does exacerbate extinctions in the microbiota after perturbation (40). The industrialized human microbiota also bears the hallmarks of sanitation, showing greater interindividual differences in microbiota composition (an indication of less microbe sharing between people) compared with traditional human populations in Papua, New Guinea, where individuals share more bacterial species with one another (20). Risking increased infectious diseases by reducing standards of sanitation would be misguided, but a better understanding of how hygienic practices shape our microbiota and the resulting impact on human health is needed. Restoring the consumption of non–disease-causing microbes may ameliorate diseases that are common among populations that consume sterile food and water (41).

Antibiotics and sanitation are intended to limit exposure to pathogenic microbes, but other practices such as the Western diet and C-section births that are not targeted at microbe control may nevertheless be having a profound effect on the microbiota.

Diet is a major driver of the composition and metabolic output of the microbiota (42–44). Humans have shifted from a diet of exclusively wild animals and gathered foods to one of domesticated livestock and agricultural produce (10,000 to 20,000 years ago) to a more recent shift to industrially produced foods, including chemically managed livestock and produce and sterilized, ultraprocessed foods containing preservatives and additives (45, 46). These shifts have resulted in a food supply capable of supporting a growing human population, but perhaps at the cost of the population's health (47).

One notable change to foodstuffs is the unintentional depletion of a major form of sustenance for the microbiota: microbiota-accessible carbohydrates (MACs; the complex carbohydrates found in the dietary fiber of edible plants such as legumes, whole grains, vegetables, nuts, etc.) (42). A high-MAC diet was commonplace when humans exclusively foraged for nutrition, and low-MAC diets have been associated with lower microbiota diversity and poor markers of health in humans and in animal models (48–50). The paucity of MACs in the industrialized diet was compensated for by additional protein, simple carbohydrates, and fat, which had the effect of altering the composition and functional output of the microbiota (43, 51). The use of additives such as emulsifiers and non-nutritive sweeteners is pervasive in industrialized food. Both have been shown to alter microbiota composition and promote intestinal inflammation. In addition, emulsifiers promote adiposity and non-nutritive sweeteners alter the metabolic output of the microbiota toward one that resembles that of type 2 diabetics (21, 52).

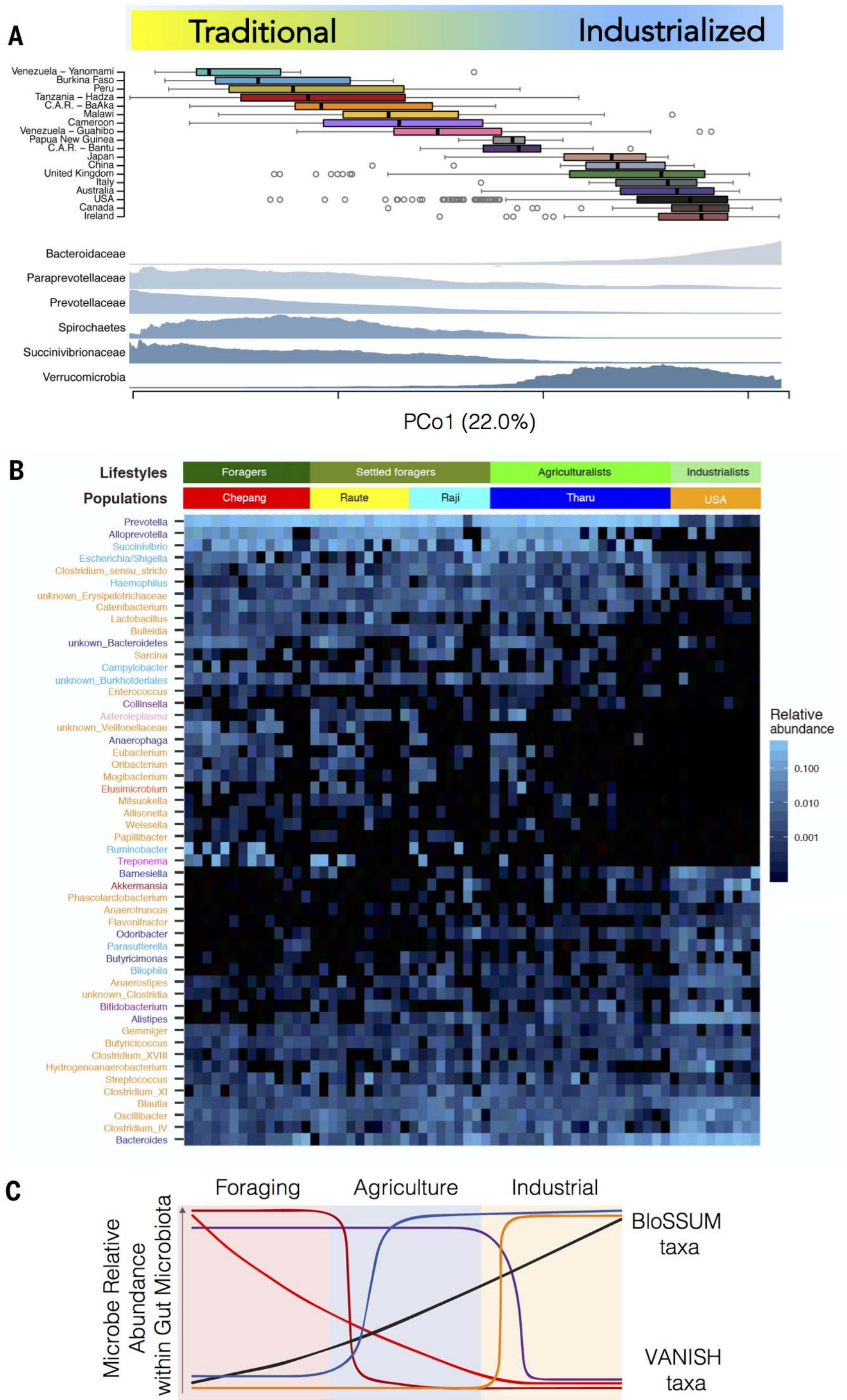


Fig. 1. The gut microbiota mirrors lifestyle across traditional and industrial populations. (**A**) Aggregation of gut microbiota composition from multiple studies separated by principal component analysis of Bray–Curtis dissimilarity of 16*S* rRNA enumerations [adapted from Smits *et al.* (*33*)]. Top panel: The first principal component explains 22% of the variation in the data from 18 populations living lifestyles spanning from uncontacted Amerindians in Venezuela (top) to fully industrialized populations in Australia, the United States, Canada, and Ireland (bottom). Bottom panel: Mapping the relative abundance of bacterial families on PCo1 reveals global patterns in the VANISH taxa, which are associated negatively with industrialized societies, and BloSSUM taxa (bloom or selected in societies of urbanization/modernization), such as the Bacteroidaceae and Verrucomicrobia. (**B**) Heat map adapted from Jha *et al*. (*31*) displaying taxa that change across lifestyles in one geographic location (Nepal) of individuals living as foragers (Chepang), settled foragers (Raute, Raji), or agriculturalists (Tharu) versus industrialized individuals in the United States. (**C**) Model adapted from Jha *et al*. (*31*) of strain loss and/or reduction versus gain and/or increase across a lifestyle gradient. Different patterns of changing abundance correspond with specific aspects of lifestyle that change as populations move away from foraging and toward urbanization. The model could also reflect the historical progression of industrialized humans from foraging (*Homo sapiens* arose ~200,000 to 300,000 years ago) to agriculture (starting 10,000 to 20,000 years ago) to industrialization (starting 100 to 200 years ago).

Small changes to the microbiota have the capacity to amplify over generations. For example, mice fed a low-MAC diet showed reduced microbiota diversity that compounded over generations. Restoration of a high-MAC diet was not sufficient to regain microbiota diversity, which indicated that species within the microbiota had gone extinct during the four-generation length of the experiment (*50*). In another study, antibiotic treatment of pregnant mice altered the microbiota of the offspring and resulted in metabolic derangement that predisposed the pups to diet-induced obesity (*53*). Similarly, C-section delivery in humans results in colonization of the infant with microbes derived from skin instead of the mother's vaginal microbiota (*54*). Acute perturbations from diet, antibiotics, and medical practices could have been propagated over generations and synergized with heightened hygiene and sanitation to result in the population-wide ecosystem reconfigurations observed today. The effects of other factors associated with an industrialized lifestyle on the microbiota, including increased sedentary behavior, stress, exposure to new chemicals (e.g., plastics, herbicides, and pesticides), and social isolation, have only begun to be explored (*55*–*57*).

Microbiota change: Good, bad, or neither?

It is not a given that the microbiota found in traditional populations, which likely shares

more commonality with that of our ancient ancestors, would improve the health of a person living in an industrialized society (4). For example, several members of a traditional gut microbiota, such as parasites, are frank pathogens. Some functions of a traditional microbiota may have beneficial effects in the context of a traditional lifestyle but may not in a more urbanized context. We have simplified these points and recognize that some parasites may confer "benefits" to human health, but how benefit is defined may depend on context and the individual. For example, parasites that protect against intestinal inflammatory diseases may cause opportunistic infections in immunocompromised individuals (58).

While remaining agnostic about broad connections between change in the microbiota and human health, it is worth considering underlying evolutionary principles that might predict whether microbiota changes are likely to be beneficial, deleterious, or neutral. A very conservative view is that until we have a good understanding of which microbes or communities are beneficial or deleterious, including how context determines this answer, we should recognize that (i) our resident microbes have the potential to affect our health in profound ways and (ii) individual lifestyle and/or medical choices and population-level lifestyle, medical, and dietary choices can change these communities. Similar to early, albeit insufficient, steps to address climate change before the full scope of the problem was understood, such as developing renewable alternatives to fossil fuels, a hedge against potential catastrophe seems warranted. In the case of our gut microbes, acting to minimize unintended loss of biodiversity is likely a wise strategy until we know more. We discuss possible strategies below.

An important question is whether loss or reduction of resident, codiversified microbes and associated functions could have a negative health impact on humans. Some properties of the human microbiota appear to have been stable during much of human evolution before industrialization. It is expected that the combined biology and genome of the human body and its commensal microorganisms would have coevolved to maximize human reproductive success (fitness) during that time (59). Because industrialized humans are interested in a long, healthy life, it is worth asking whether long life is consistent with the reproductive success of early humans. The reproductive success of modern hunter-gatherers corresponds to being long lived (as demonstrated by evidence supporting the patriarch hypothesis); therefore, the components of the microbiome that lived within humans throughout most of our existence as a species likely promote biology consistent with a long, healthy life (60).

From the microbial point of view, a bacterial species is chiefly concerned with making more of itself. Therefore, it is worth considering whether it is possible for members of the microbiota that increase host health and longevity to arise. In other words, the question is not only whether the interests of host and microbiota are aligned (i.e., to promote a long, healthy life of the host), but whether microbes that promote the health and longevity of their hosts are retained and favored over evolutionary time.

Gut-resident microbes that improve host health and life span are most likely to arise when the health-promoting function does not incur a short-term fitness cost to themselves (61, 62). For example, imagine a microbial pathway that not only generates energy for the microbe by fermenting a dietary complex carbohydrate but also produces a fermentation end product that can be absorbed by the host and play beneficial metabolic and/or regulatory roles. These microbes would contribute to host health without incurring a fitness cost and could be selected over time as a result of host fitness, longevity, and transmission to offspring and other individuals. We might expect that loss of these coevolved microbes and associated functions would have a negative health impact.

The industrialized microbiota could be considered better adapted to an industrialized host lifestyle by harboring more resistance to antibiotics and being less proficient at dietary fiber degradation. However, such a microbiota may not be optimized for our health.

Strategies to protect and, when appropriate, restore the microbiome

Learning how to minimize harm to an ecosystem is an easier prospect than rebuilding one that has deteriorated; however, the realization of an ecosystem's importance often only becomes apparent after major change has taken place. In the case of the gut microbiota, we may have to confront the daunting task of reconfiguring an ecosystem that we are just beginning to understand. Biodiverse ecosystems are characterized by complex networks of interactions; delete or add one node and the cascade of changes through the network of interactions can be difficult to anticipate. Predicting ecosystem changes from species reintroduction, such as wolves into Yellowstone National Park, is a challenge long faced by conservation biologists (63, 64) (Fig. 2A).

Fecal microbiota transplantation (FMT) is an example of how ecosystem remodeling through multispecies rewilding can be applied to the gut microbiota. In this procedure, all of the bacterial species of a healthy human donor's stool microbiota are introduced into a diseased recipient in an attempt to reconfigure a maladaptive ecosystem (Fig. 2B) (65). FMT has been highly effective in treating *Clostridium difficile* infection (CDI) refractory to conventional antibiotic-based treatment (66). Although this procedure cures CDI, the addition of hundreds of microbial species into an equally complex, although disrupted, ecosystem results in an unpredictable community that is composed of strains from the donor,

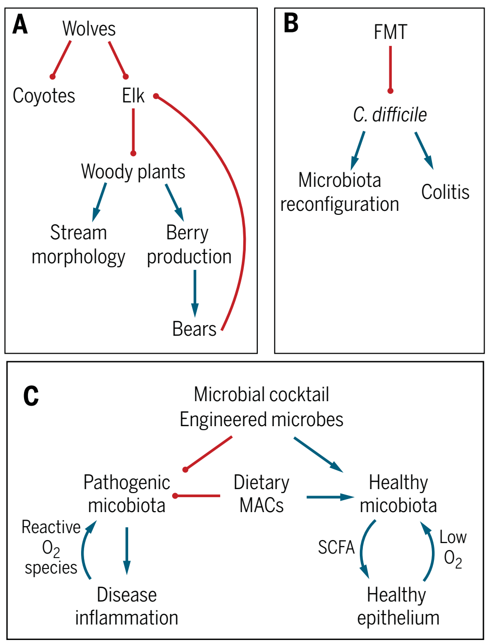


Fig. 2. Interaction networks in Yellowstone and the gut microbiota. (**A**) Gray wolves were introduced into Yellowstone National Park in 1995 to control the swelling elk population (105). The rewilding of Yellowstone set off a trophic cascade that resulted in a decreasing elk population (thereby promoting new growth in aspens), an increase in berries available to bears, and stream morphology changes caused by increased woody plants (64). This provides an example of how wildlife management can be used to restore a more diverse and perhaps functional ecosystem, as well as how reintroduction of species into a habitat can lead to unanticipated changes to that ecosystem. (**B**) Rewilding of a *C. difficile*–infected microbiota by FMT results in largely predictable outcomes in host health, but the specifics of the resulting microbiota composition are difficult to predict. (**C**) Long-term strategies for managing the microbiota include precision approaches of adding defined cocktails of microbes, engineered bacterial species, and improving ecosystem habitat quality. For example, increasing dietary MACs encourages commensal growth and provides fermentation end products such as butyrate to the epithelium, which can help keep oxygen tensions lower in the gut and prevent the growth of facultative anaerobes with pathogenic potential (106).

recipient, and other sources (67, 68). CDI represents an extreme case of ecosystem disruption; therefore, the lack of precision in dictating the resulting community after ecosystem rewilding is clinically tolerable, as almost any resulting microbiota configuration lacking or minimizing *C. difficile* is preferred. However, FMTs are not an ideal long-term solution for the treatment of many diseases. In many cases, they are simply ineffective, and in others, the unintended consequences may include transmission of antibiotic-resistant microbes or other infectious agents and the transference of unwanted phenotypes from the donor (69). Gut microbiota rewilding through FMT has currently only been consistently successful for *C. difficile* cases. Similar to cases of animal reintroduction in macroecosystems, success as defined by the ability of these reintroduced species to thrive has been mixed (70).

Targeted rewilding through discrete changes in habitat quality or the introduction of specific species chosen based on known interactions may be a more predictable and successful approach to ecosystem management in a disrupted gut microbiota. Habitat quality is a key element of success in macroecosystem restoration and is also an important consideration in the gut (71). Ecosystems are made up of interacting species and their physicochemical environment. Factors that influence the suitability of the gut habitat, including temperature, pH, osmolality, redox status, water activity, and chemical and nutrient availability, will likely affect the success of microbiota reconfiguration efforts. Mice chronically infected with *C. difficile* can be effectively treated using a diet containing MACs. This simple change to habitat quality enabled the recovery of a robust indigenous community and reestablished important functions such as short-chain fatty acid (SCFA) production (72). Diet can also create a niche for a newly introduced microbial strain to colonize. For instance, feeding mice the seaweed polysaccharide porphyran allowed engraftment of a porphyran-utilizing Bacteroides strain (73). This example of engrafting a new species into a microbiota may provide a strategy that can be extended to help targeted rewilding (Fig. 2C).

An additional challenge to managing ecosystems is identifying the features within an ecosystem that are "beneficial" and thus worthy of conservation. One strategy used by ecologists is to assess the "services" provided by an ecosystem. Ecosystem services, popularized in the Millennium Ecosystem Assessment, enable value to be placed on different components of an ecosystem (74). For example, if a lake provides fresh drinking water and recreation (swimming, fishing), then pollution of that lake would put those services in jeopardy. Likewise, we can consider the ecosystem services supplied by the gut microbiota (75) (Fig. 3).

However, determining whether a microbiota ecosystem service is beneficial is difficult enough in itself, and establishing whether this benefit is universal or specific to a subpopulation of people or even only one individual, a developmental period of life, or during disease or reproduction adds complexity. For example, extraction of calories was an important microbiota ecosystem service rendered in the preindustrialized world, but when eating modern, calorie-dense foods, this service becomes less important.

Studying microbiota configurations in different contexts may reveal associations that are positive for human health. For example, work on the gut microbiota in individuals undergoing immunotherapy to treat cancer has shown associations between specific microbiota components and improved outcomes (76). Although many specifics remain to be determined, these findings are consistent with the ability of different microbiotas and their services, such as SCFA production, to alter host immune status and function. Unfortunately, such observational work is usually conducted on people living in industrialized countries and therefore is limited in the microbiota configurations and features that are queried.

If humans have developed a dependence upon microbiota services that have been lost during industrialization, then might reintroduction of these services be analogous to complementing a lost portion of human biology and provide broad benefit? Even if this is not the case, given the recent success of prophylactic antibiotics in low- and middle-income countries in improving health and reducing mortality in children, rewilding the microbiota after treatment using defined key strains may become a standard treatment to aid in ecosystem recovery (77). Should this be the case, then considerations of how to make reintroductions self-sustaining, especially in the face of spreading industrialization, will be important.

Toward a more sustainable relationship with our microbiota

The goals of a managed microbiota should be to optimize ecosystem services to prevent disease and improve health and longevity. Optimization requires precise, targeted approaches that consider an individual's genotype, microbiome, or subcategory of disease. Given the large global health impact, strategies to protect the microbiome in all populations should

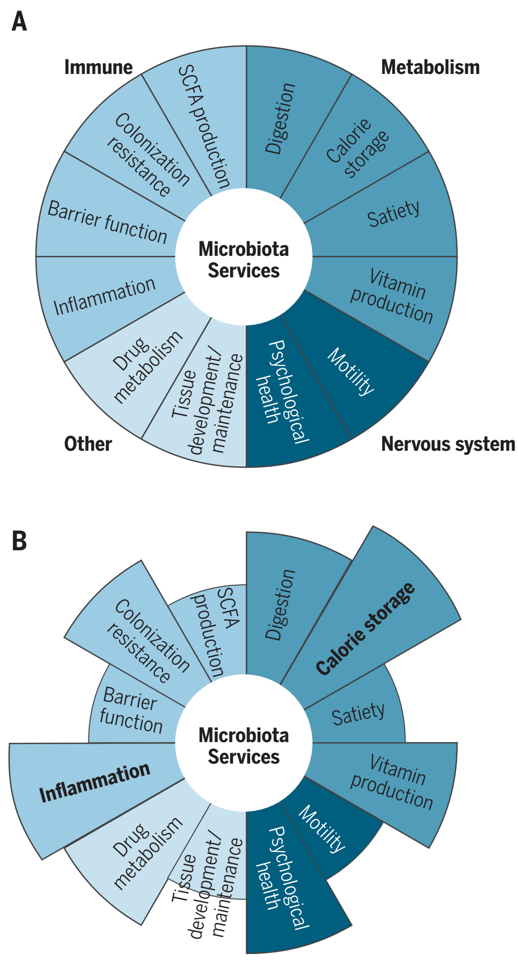


Fig. 3. Ecosystem services of the gut microbiota. Identifying the benefits provided by the gut microbiome to human health is one way to determine when the ecosystem is functioning well. (**A**) List of benefits provided by the gut microbiota. This list is not intended to be comprehensive, and the categorization is only one of many possibilities, but it is presented as a potentially useful framework for conceptualizing how to value specific features of microbiota. (**B**) Current data suggest that, along with the shift in the composition of the industrialized microbiota, certain services may be lost or out of balance, resulting in suboptimal states of host physiology or disease. A more nuanced understanding of which services are beneficial and in what context will be enabled by longitudinal high-dimensional profiling of microbiome and host biology combined with long-term monitoring of health in humans.

Box. Sustainable ecosystem management approaches.

As we continue to learn of the multitude of benefits afforded by our microbial symbionts, developing alternative strategies to manage microbial ecosystems will enable us to promote short- and long-term public health priorities simultaneously (87). Listed here are a few examples of successes in using beneficial microbes to manage microbial ecosystems.

Wound healing

Sterility in skin-injury repair has been viewed as an important factor in effective wound healing. However, maintaining a sterile wound-healing environment is a difficult prospect considering the exposure of most wounds to the environment (88). Recent evidence suggests that populating wounds with commensal microbes can reduce infections after surgery and minimize the need for antibiotic treatment (89). Similar strategies are also being tested in treating skin conditions including atopic dermatitis (clinical trial NCT03018275) and acute wounds (90).

Probiotics in hospital cleaning

Health care–associated infections are pervasive in both high- and low-income countries and are a leading cause of death in the United States (91). Germicidal treatments of hospital surfaces are not completely effective, leaving behind dangerous pathogens, some of which can inhabit surfaces for months and also lead to increasing antibiotic resistance. The use of probiotic-containing cleaners can be an effective, alternative method to decontaminate hospital surfaces that does not select for antibiotic-resistant strains (92).

Food safety

Concerns over increasing antibiotic resistance, consumption of antibiotic-laden meat, and antibiotic-induced reduction of natural resistance to pathogens have led to the exploration of alternatives to antibiotics in livestock. Probiotic use in chickens has resulted in better growth rates, reductions in pathogen load and antibiotic resistance genes, and improved egg quality (93, 94). Probiotics have also been used to prevent infections and improve milk production in dairy cows and to aid growth in beef cattle (95). Use of probiotics is also beneficial in aquaculture, improving water quality, resistance to pathogens, and growth (96).

Fighting infections in humans

There is growing evidence that the use of beneficial bacteria is a promising path forward for managing pathogenic microbes in humans (97). Probiotics can reduce the duration and severity of infectious diarrhea and may be an effective alternative to antibiotics in the treatment and prevention of bacterial vaginosis (98, 99). A synbiotic mixture of *Lactobacillus plantarum* and fructo-oligosaccharides reduced the incidence of sepsis and lowered rates of respiratory tract infection in a cohort of infants from rural India (100). The use of bacteriophage to control pathogens, especially those that are resistant to multiple antibiotics, is another emerging alternative with recent success (101).

Priming an ecosystem to fight cancer

Antibiotics are commonly used in cancer treatment to minimize the risk of infection in a patient population with a disrupted immune system. However, in animal models, antibiotic treatment can alter the microbiota in ways that reduce treatment efficacy (102, 103). In fact, specific manipulation of the microbiota improved immunotherapy-based tumor control in a mouse model of melanoma (102, 103). Optimization of the microbiota to optimize immune status, whether augmenting immunotherapy or enabling bone marrow transplantation, will likely be integral to the future treatment of diseases such as cancer.

C-section delivery

Given newly acquired data about the importance of early microbiota assembly in the health of the infant, a rethinking of medicalized birth is warranted. A recent pilot study showed that infants delivered by C-section who were seeded with their mother's vaginal microbes developed microbiota more closely resembling those of vaginally delivered infants (104). Future studies are required to determine whether vaginal seeding after C-section delivery provides any lifelong health benefit to the infant.

be considered to maximize the palette of microbial and molecular tools available. Efforts are under way to archive the microbial diversity found in the gut of humans around the globe (35, 36). Whether these efforts will result in new therapeutics remains to be seen, but at the very least they provide a time capsule of microbial diversity in a rapidly industrializing world. Industrialization of the microbiome, and its accompanying loss or reduction of certain species, can occur on a time scale of months within an individual, creating some urgency for the banking of vulnerable species (78). An additional challenge is navigating the changing restrictions on the distribution of bacterial strains for research and therapeutic development while protecting the rights and recognizing the contribution of the people from which they came (79, 80).

Reshaping ingrained aspects of industrialized societies to moderate practices that have negative impacts on the microbiota will be a challenge but will be more practical than reversion to preindustrial practices (see Box: Sustainable ecosystem management approaches). Antibiotic use will remain an important aspect of industrial life; however, regulation in clinical and agricultural settings is needed to maintain efficacy and to protect the microbiome. Similarly, rationally engineered microbial "cocktails" or fermented foods could offer safe microbe exposure to compensate for sanitization. Government subsidies similar to those provided for certain crops could be justified to make MAC-rich and fermented foods cheaper and more widely available. Until food policy reflects the findings of biomedical research, short-term solutions, such as supplementing processed foods with MACs or probiotic bacteria, could provide a gradual progression toward health-optimizing food systems in industrialized countries.

Expanding cohort and interventional studies in humans from a wide representation of humans while simultaneously documenting microbiome and health changes is key for healthy, sustainable microbiota. Numerous associations have been made between the microbiota and human disease, but additional microbiome datasets from longitudinal, prospective observational and interventional studies of humans will provide insight into causal relationships. High-resolution measurements of host biology, including "omics" approaches and high-dimensional immune profiling, will be important to elucidate the specific lifestyle practices that lead to the most meaningful microbiome changes for human health (44, 81, 82). Animal models informed by human-derived data can be used to perform controlled studies with the goal of developing strategies to rebuild and maintain a healthy microbiota (83).

A connection with the greater ecosystem

Some of the specific forces that are bad for Earth appear also to harm our microbiota. For example, animal meat production removes forest habitat for pasture and results in increased methane production. Excessive meat consumption has been coupled to trimethylamine-*N*-oxide (TMAO) production by the microbiota, and TMAO is a risk factor for cardiovascular events (84). It may be wise to approach climate and health and microbiota sustainability simultaneously to identify solutions that align

Earth and human health (i.e., One World, One Health) (*85*). Given that environmentally sustainable agricultural practices are compatible with producing food generally recognized to promote health, solutions for the planet and human health may be compatible (*86*). As Earth's microbes adapt to our changing environment, we can expect our body's ecosystem to reflect our external environment in ways that are difficult to anticipate. Determining microbial or molecular equivalents of rewilding will require a much better understanding of community dynamics and host–microbiota interactions than we presently have. Continually monitoring and managing a healthy internal ecosystem may be an effective strategy to combat and prevent the litany of chronic diseases that are currently spreading with industrialization.

REFERENCES AND NOTES

1. C. J. Stewart *et al*., Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018). doi: 10.1038/s41586-018-0617-x; pmid: 30356187
2. N. A. Bokulich *et al*., Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* **8**, 343ra82 (2016). doi: 10.1126/scitranslmed.aad7121; pmid: 27306664
3. A. H. Moeller *et al*., Cospeciation of gut microbiota with hominids. *Science* **353**, 380–382 (2016). doi: 10.1126/science.aaf3951; pmid: 27463672
4. E. D. Sonnenburg, J. L. Sonnenburg, The ancestral and industrialized gut microbiota and implications for human health. *Nat. Rev. Microbiol.* **17**, 383–390 (2019). doi: 10.1038/s41579-019-0191-8; pmid: 31089293
5. Y. Belkaid, O. J. Harrison, Homeostatic Immunity and the Microbiota. *Immunity* **46**, 562–576 (2017). doi: 10.1016/j.immuni.2017.04.008; pmid: 28423337
6. B. O. Schroeder, F. Bäckhed, Signals from the gut microbiota to distant organs in physiology and disease. *Nat. Med.* **22**, 1079–1089 (2016). doi: 10.1038/nm.4185; pmid: 27711063
7. A. Sarkar *et al*., The Microbiome in Psychology and Cognitive Neuroscience. *Trends Cogn. Sci.* **22**, 611–636 (2018). doi: 10.1016/j.tics.2018.04.006; pmid: 29907531
8. G. Sharon, T. R. Sampson, D. H. Geschwind, S. K. Mazmanian, The central nervous system and the gut microbiome. *Cell* **167**, 915–932 (2016). doi: 10.1016/j.cell.2016.10.027; pmid: 27814521
9. J. L. Sonnenburg, F. Bäckhed, Diet-microbiota interactions as moderators of human metabolism. *Nature* **535**, 56–64 (2016). doi: 10.1038/nature18846; pmid: 27383980
10. T. Gensollen, S. S. Iyer, D. L. Kasper, R. S. Blumberg, How colonization by microbiota in early life shapes the immune system. *Science* **352**, 539–544 (2016). doi: 10.1126/science.aad9378; pmid: 27126036
11. J. Walter, R. Ley, The human gut microbiome: Ecology and recent evolutionary changes. *Annu. Rev. Microbiol.* **65**, 411–429 (2011). doi: 10.1146/annurev-micro-090110-102830; pmid: 21682646
12. E. K. Costello, K. Stagaman, L. Dethlefsen, B. J. Bohannan, D. A. Relman, The application of ecological theory toward an understanding of the human microbiome. *Science* **336**, 1255–1262 (2012). doi: 10.1126/science.1224203; pmid: 22674335
13. L. Dethlefsen, D. A. Relman, Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. U.S.A.* **108** (Suppl 1), 4554–4561 (2011). doi: 10.1073/pnas.1000087107; pmid: 20847294
14. M. J. Blaser, S. Falkow, What are the consequences of the disappearing human microbiota? *Nat. Rev. Microbiol.* **7**, 887–894 (2009). doi: 10.1038/nrmicro2245; pmid: 19898491
15. Centers for Disease Control and Prevention, Achievements in Public Health, 1900-1999: Control of Infectious Diseases. *MMWR* **48**, 621–629 (1999).
16. I. Roca *et al*., The global threat of antimicrobial resistance: Science for intervention. *New Microbes New Infect.* **6**, 22–29 (2015). doi: 10.1016/j.nmni.2015.02.007; pmid: 26029375
17. E. Pasolli *et al*., Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* **176**, 649–662.e20 (2019). doi: 10.1016/j.cell.2019.01.001; pmid: 30661755
18. S. Nayfach, Z. J. Shi, R. Seshadri, K. S. Pollard, N. C. Kyrpides, New insights from uncultivated genomes of the global human gut microbiome. *Nature* **568**, 505–510 (2019). doi: 10.1038/s41586-019-1058-x; pmid: 30867587
19. A. Almeida *et al*., A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499–504 (2019). doi: 10.1038/s41586-019-0965-1; pmid: 30745586
20. I. Martínez *et al*., The gut microbiota of rural papua new guineans: Composition, diversity patterns, and ecological processes. *Cell Rep.* **11**, 527–538 (2015). doi: 10.1016/j.celrep.2015.03.049; pmid: 25892234
21. J. Suez *et al*., Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* **514**, 181–186 (2014). doi: 10.1038/nature13793; pmid: 25231862
22. N. T. Mueller, E. Bakacs, J. Combellick, Z. Grigoryan, M. G. Dominguez-Bello, The infant microbiome development: Mom matters. *Trends Mol. Med.* **21**, 109–117 (2015). doi: 10.1016/j.molmed.2014.12.002; pmid: 25578246
23. S. J. O'Keefe *et al*., Fat, fibre and cancer risk in African Americans and rural Africans. *Nat. Commun.* **6**, 6342 (2015). doi: 10.1038/ncomms7342; pmid: 25919227
24. M. J. Blaser, The past and future biology of the human microbiome in an age of extinctions. *Cell* **172**, 1173–1177 (2018). doi: 10.1016/j.cell.2018.02.040; pmid: 29522739
25. GBD 2015 Mortality and Causes of Death Collaborators, Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1459–1544 (2016). doi: 10.1016/S0140-6736(16)31012-1; pmid: 27733281
26. R. Lozano *et al*., Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2095–2128 (2012). doi: 10.1016/S0140-6736(12)61728-0; pmid: 23245604
27. P. M. Vitousek, H. A. Mooney, J. Lubchenco, J. M. Melillo, Human Domination of Earth's Ecosystems. *Science* **277**, 494–499 (1997). doi: 10.1126/science.277.5325.494
28. G. R. Walther *et al*., Ecological responses to recent climate change. *Nature* **416**, 389–395 (2002). doi: 10.1038/416389a; pmid: 11919621
29. I. M. Maclean, R. J. Wilson, Recent ecological responses to climate change support predictions of high extinction risk. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12337–12342 (2011). doi: 10.1073/pnas.1017352108; pmid: 21746924
30. G. K. Fragiadakis *et al*., Links between environment, diet, and the hunter-gatherer microbiome. *Gut Microbes* **10**, 216–227 (2019). doi: 10.1080/19490976.2018.1494103; pmid: 30118385
31. A. R. Jha *et al*., Gut microbiome transition across a lifestyle gradient in Himalaya. *PLOS Biol.* **16**, e2005396 (2018). doi: 10.1371/journal.pbio.2005396; pmid: 30439937
32. P. Vangay *et al*., US Immigration Westernizes the Human Gut Microbiome. *Cell* **175**, 962–972.e10 (2018). doi: 10.1016/j.cell.2018.10.029; pmid: 30388453
33. S. A. Smits *et al*., Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* **357**, 802–806 (2017). doi: 10.1126/science.aan4834; pmid: 28839072
34. J. L. Sonnenburg *et al*., Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science* **307**, 1955–1959 (2005). doi: 10.1126/science.1109051; pmid: 15790854
35. M. G. Dominguez-Bello, R. Knight, J. A. Gilbert, M. J. Blaser, Preserving microbial diversity. *Science* **362**, 33–34 (2018). doi: 10.1126/science.aau8816; pmid: 30287652
36. T. Rabesandratana, Microbiome conservancy stores global fecal samples. *Science* **362**, 510–511 (2018). doi: 10.1126/science.362.6414.510; pmid: 30385553
37. M. J. Blaser, Antibiotic use and its consequences for the normal microbiome. *Science* **352**, 544–545 (2016). doi: 10.1126/science.aad9358; pmid: 27126037
38. A. Schulfer, M. J. Blaser, Risks of antibiotic exposures early in life on the developing microbiome. *PLOS Pathog.* **11**, e1004903 (2015). doi: 10.1371/journal.ppat.1004903; pmid: 26135581
39. G. Campbell-Platt, Fermented foods: A world perspective. *Food Res. Int.* **27**, 253–257 (1994). doi: 10.1016/0963-9969(94)90093-0
40. C. Tropini *et al*., Transient osmotic perturbation causes long-term alteration to the gut microbiota. *Cell* **173**, 1742–1754 (2018).
41. S. N. Chilton, J. P. Burton, G. Reid, Inclusion of fermented foods in food guides around the world. *Nutrients* **7**, 390–404 (2015). doi: 10.3390/nu7010390; pmid: 25580813
42. E. D. Sonnenburg, J. L. Sonnenburg, Starving our microbial self: The deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab.* **20**, 779–786 (2014). doi: 10.1016/j.cmet.2014.07.003; pmid: 25156449
43. L. A. David *et al*., Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014). doi: 10.1038/nature12820; pmid: 24336217
44. G. D. Wu *et al*., Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011). doi: 10.1126/science.1208344; pmid: 21885731
45. A. Snir *et al*., The origin of cultivation and proto-weeds, long before neolithic farming. *PLOS ONE* **10**, e0131422 (2015). doi: 10.1371/journal.pone.0131422; pmid: 26200895
46. C. A. Monteiro, J. C. Moubarac, G. Cannon, S. W. Ng, B. Popkin, Ultra-processed products are becoming dominant in the global food system. *Obes. Rev.* **14** (suppl. 2), 21–28 (2013). doi: 10.1111/obr.12107; pmid: 24102801
47. M. B. Cole, M. A. Augustin, M. J. Robertson, J. M. Manners, The science of food security. *NPJ Sci. Food* **2**, 14 (2018).
48. S. B. Eaton, The ancestral human diet: What was it and should it be a paradigm for contemporary nutrition? *Proc. Nutr. Soc.* **65**, 1–6 (2006). doi: 10.1079/PNS2005471; pmid: 16441938
49. A. Cotillard *et al*., Dietary intervention impact on gut microbial gene richness. *Nature* **500**, 585–588 (2013). doi: 10.1038/nature12480; pmid: 23985875
50. E. D. Sonnenburg *et al*., Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**, 212–215 (2016). doi: 10.1038/nature16504; pmid: 26762459
51. S. Devkota *et al*., Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10-/- mice. *Nature* **487**, 104–108 (2012). doi: 10.1038/nature11225; pmid: 22722865
52. B. Chassaing *et al*., Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature* **519**, 92–96 (2015). doi: 10.1038/nature14232; pmid: 25731162
53. L. M. Cox *et al*., Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell* **158**, 705–721 (2014). doi: 10.1016/j.cell.2014.05.052; pmid: 25126780
54. M. G. Dominguez-Bello *et al*., Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11971–11975 (2010). doi: 10.1073/pnas.1002601107; pmid: 20566857
55. C. Bressa *et al*., Differences in gut microbiota profile between women with active lifestyle and sedentary women. *PLOS ONE* **12**, e0171352 (2017). doi: 10.1371/journal.pone.0171352; pmid: 28187199
56. R. T. Liu, The microbiome as a novel paradigm in studying stress and mental health. *Am. Psychol.* **72**, 655–667 (2017). doi: 10.1037/amp0000058; pmid: 29016169
57. K. A. Dill-McFarland *et al*., Close social relationships correlate with human gut microbiota composition. *Sci. Rep.* **9**, 703 (2019). doi: 10.1038/s41598-018-37298-9; pmid: 30679677
58. J. V. Weinstock, Do we need worms to promote immune health? *Clin. Rev. Allergy Immunol.* **49**, 227–231 (2015). doi: 10.1007/s12016-014-8458-3; pmid: 25326880
59. C. De Mazancourt, M. Loreau, U. Dieckmann, Understanding mutualism when there is adaptation to the partner. *J. Ecol.* **93**, 305–314 (2005). doi: 10.1111/j.0022-0477.2004.00952.x
60. F. Marlowe, The patriarch hypothesis: An alternative explanation of menopause. *Hum. Nat.* **11**, 27–42 (2000). doi: 10.1007/s12110-000-1001-7; pmid: 26193094
61. K. R. Foster, J. Schluter, K. Z. Coyte, S. Rakoff-Nahoum, The evolution of the host microbiome as an ecosystem on a leash. *Nature* **548**, 43–51 (2017). doi: 10.1038/nature23292; pmid: 28770836
62. S. Rakoff-Nahoum, K. R. Foster, L. E. Comstock, The evolution of cooperation within the gut microbiota.

Nature **533**, 255–259 (2016). doi: 10.1038/nature17626; pmid: 27111508

63. J. Lorimer *et al*., Rewilding: Science, practice, and politics. *Annu. Rev. Environ. Resour.* **40**, 39–62 (2015). doi: 10.1146/annurev-environ-102014-021406
64. W. J. Ripple *et al*., Status and ecological effects of the world's largest carnivores. *Science* **343**, 1241484 (2014). doi: 10.1126/science.1241484; pmid: 24408439
65. A. Khoruts, M. J. Sadowsky, Understanding the mechanisms of faecal microbiota transplantation. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 508–516 (2016). doi: 10.1038/nrgastro.2016.98; pmid: 27329806
66. E. van Nood *et al*., Duodenal infusion of donor feces for recurrent *Clostridium difficile*. *N. Engl. J. Med.* **368**, 407–415 (2013). doi: 10.1056/NEJMoa1205037; pmid: 23323867
67. C. Staley, C. R. Kelly, L. J. Brandt, A. Khoruts, M. J. Sadowsky, Complete microbiota engraftment is not essential for recovery from recurrent *Clostridium difficile* infection following fecal microbiota transplantation. *mBio* **7**, e01965-16 (2016). doi: 10.1128/mBio.01965-16; pmid: 27999162
68. C. S. Smillie *et al*., Strain tracking reveals the determinants of bacterial engraftment in the human gut following fecal microbiota transplantation. *Cell Host Microbe* **23**, 229–240.e5 (2018). doi: 10.1016/j.chom.2018.01.003; pmid: 29447696
69. N. Alang, C. R. Kelly, Weight gain after fecal microbiota transplantation. *Open Forum Infect. Dis.* **2**, ofv004 (2015). doi: 10.1093/ofid/ofv004; pmid: 26034755
70. P. J. Seddon, C. J. Griffiths, P. S. Soorae, D. P. Armstrong, Reversing defaunation: Restoring species in a changing world. *Science* **345**, 406–412 (2014). doi: 10.1126/science.1251818; pmid: 25061203
71. B. Griffith, J. M. Scott, J. W. Carpenter, C. Reed, Translocation as a species conservation tool: Status and strategy. *Science* **245**, 477–480 (1989). doi: 10.1126/science.245.4917.477; pmid: 17750257
72. A. J. Hryckowian *et al*., Microbiota-accessible carbohydrates suppress *Clostridium difficile* infection in a murine model. *Nat. Microbiol.* **3**, 662–669 (2018). doi: 10.1038/s41564-018-0150-6; pmid: 29686297
73. E. S. Shepherd, W. C. DeLoache, K. M. Pruss, W. R. Whitaker, J. L. Sonnenburg, An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* **557**, 434–438 (2018). doi: 10.1038/s41586-018-0092-4; pmid: 29743671
74. M. E. Assesment, *Ecosystems and Human Well-Being: A Framework for Assessment* (Island, 2003).
75. E. A. McKenney, K. Koelle, R. R. Dunn, A. D. Yoder, The ecosystem services of animal microbiomes. *Mol. Ecol.* **27**, 2164–2172 (2018). doi: 10.1111/mec.14532; pmid: 29427300
76. S. Roy, G. Trinchieri, Microbiota: A key orchestrator of cancer therapy. *Nat. Rev. Cancer* **17**, 271–285 (2017). doi: 10.1038/nrc.2017.13; pmid: 28303904
77. J. D. Keenan *et al*., Azithromycin to reduce childhood mortality in Sub-Saharan Africa. *N. Engl. J. Med.* **378**, 1583–1592 (2018). doi: 10.1056/NEJMoa1715474; pmid: 29694816
78. J. Sonnenburg, E. Sonnenburg, A microbiota assimilation. *Cell Metab.* **28**, 675–677 (2018). doi: 10.1016/j.cmet.2018.10.010; pmid: 30403987
79. J. Overmann, A. H. Scholz, Microbiological research under the Nagoya protocol: Facts and fiction. *Trends Microbiol.* **25**, 85–88 (2017). doi: 10.1016/j.tim.2016.11.001; pmid: 27887771
80. I. N. Okeke, C. P. Babalola, D. K. Byarugaba, A. Djimde, O. R. Osoniyi, Broadening participation in the sciences within and from Africa: Purpose, challenges, and prospects. *CBE Life Sci. Educ.* **16**, es2 (2017). doi: 10.1187/cbe.15-12-0265; pmid: 28408409
81. G. D. Wu *et al*., Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. *Gut* **65**, 63–72 (2016). doi: 10.1136/gutjnl-2014-308209; pmid: 25431456
82. G. S. Abu-Ali *et al*., Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat. Microbiol.* **3**, 356–366 (2018). doi: 10.1038/s41564-017-0084-4; pmid: 29335555
83. J. L. Gehrig *et al*., Effects of microbiota-directed foods in gnotobiotic animals and undernourished children. *Science* **365**, eaau4732 (2019). doi: 10.1126/science.aau4732; pmid: 31296738
84. W. Zhu *et al*., Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk. *Cell* **165**, 111–124 (2016). doi: 10.1016/j.cell.2016.02.011; pmid: 26972052
85. P. D. van Helden, L. S. van Helden, E. G. Hoal, One world, one health. Humans, animals and the environment are inextricably linked—A fact that needs to be remembered and exploited in our modern approach to health. *EMBO Rep.* **14**, 497–501 (2013). doi: 10.1038/embor.2013.61; pmid: 23681441
86. W. Willett *et al*., Food in the Anthropocene: The EAT-Lancet Commission on healthy diets from sustainable food systems. *Lancet* **393**, 447–492 (2019). doi: 10.1016/S0140-6736(18)31788-4; pmid: 30660336
87. S. R. Modi, J. J. Collins, D. A. Relman, Antibiotics and the gut microbiota. *J. Clin. Invest.* **124**, 4212–4218 (2014). doi: 10.1172/JCI72333; pmid: 25271726
88. T. R. Johnson *et al*., The cutaneous microbiome and wounds: New molecular targets to promote wound healing. *Int. J. Mol. Sci.* **19**, 2699 (2018). doi: 10.3390/ijms19092699; pmid: 30208569
89. N. Kasatpibal *et al*., Effectiveness of probiotic, prebiotic, and synbiotic therapies in reducing postoperative complications: A systematic review and network meta-analysis. *Clin. Infect. Dis.* **64**, S153–S160 (2017). pmid: 28475793
90. J. C. Valdez, M. C. Peral, M. Rachid, M. Santana, G. Perdigon, Interference of *Lactobacillus plantarum* with *Pseudomonas aeruginosa* in vitro and in infected burns: The potential use of probiotics in wound treatment. *CMI* **11**, 472–479 (2005).
91. M. Haque, M. Sartelli, J. McKimm, M. Abu Bakar, Health care-associated infections - an overview. *Infect. Drug Resist.* **11**, 2321–2333 (2018). doi: 10.2147/IDR.S177247; pmid: 30532565
92. E. Caselli, Hygiene: Microbial strategies to reduce pathogens and drug resistance in clinical settings. *Microb. Biotechnol.* **10**, 1079–1083 (2017). doi: 10.1111/1751-7915.12755; pmid: 28677216
93. M. Alagawany *et al*., The use of probiotics as eco-friendly alternatives for antibiotics in poultry nutrition. *Environ. Sci. Pollut. Res. Int.* **25**, 10611–10618 (2018). doi: 10.1007/s11356-018-1687-x; pmid: 29532377
94. E. M. Saliu, W. Vahjen, J. Zentek, Types and prevalence of extended-spectrum beta-lactamase producing Enterobacteriaceae in poultry. *Anim. Health Res. Rev.* **18**, 46–57 (2017). doi: 10.1017/S1466252317000020; pmid: 28641596
95. Y. Uyeno, S. Shigemori, T. Shimosato, Effect of probiotics/prebiotics on cattle health and productivity. *Microbes Environ.* **30**, 126–132 (2015). doi: 10.1264/jsme2.ME14176; pmid: 26004794
96. J. L. Balcázar *et al*., The role of probiotics in aquaculture. *Vet. Microbiol.* **114**, 173–186 (2006). doi: 10.1016/j.vetmic.2006.01.009; pmid: 16490324
97. S. King *et al*., Does probiotic consumption reduce antibiotic utilization for common acute infections? A systematic review and meta-analysis. *Eur. J. Public Health* **29**, 494–499 (2019). doi: 10.1093/eurpub/cky185; pmid: 30219897
98. J. A. Applegate, C. L. Fischer Walker, R. Ambikapathi, R. E. Black, Systematic review of probiotics for the treatment of community-acquired acute diarrhea in children. *BMC Public Health* **13** (suppl. 3), S16 (2013). pmid: 24564646
99. J. M. Kim, Y. J. Park, Probiotics in the prevention and treatment of postmenopausal vaginal infections: Review article. *J. Menopausal Med.* **23**, 139–145 (2017). doi: 10.6118/jmm.2017.23.3.139; pmid: 29354612
100. P. Panigrahi *et al*., A randomized synbiotic trial to prevent sepsis among infants in rural India. *Nature* **548**, 407–412 (2017). doi: 10.1038/nature23480; pmid: 28813414
101. R. T. Schooley *et al*., Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. *Antimicrob. Agents Chemother.* **61**, e00954-17 (2017). doi: 10.1128/AAC.00954-17; pmid: 28807909
102. M. Vétizou *et al*., Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015). doi: 10.1126/science.aad1329; pmid: 26541610
103. A. Sivan *et al*., Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* **350**, 1084–1089 (2015). doi: 10.1126/science.aac4255; pmid: 26541606
104. M. G. Dominguez-Bello *et al*., Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. *Nat. Med.* **22**, 250–253 (2016). doi: 10.1038/nm.4039; pmid: 26828196
105. E. E. Bangs, S. H. Fritts, Reintroducing the gray wolf to Central Idaho and Yellowstone National Park. *Wildl. Soc. Bull.* 24, 402–413 (1996).
106. Y. Litvak, M. X. Byndloss, A. J. Bäumler, Colonocyte metabolism shapes the gut microbiota. *Science* **362**, eaat9076 (2018). doi: 10.1126/science.aat9076; pmid: 30498100

ACKNOWLEDGMENTS

We thank members of the Sonnenburg lab and collaborators for helpful discussions. **Funding:** This work was supported by the NIH (R01-DK085025 and DP1-AT00989201). J.L.S. is a Chan Zuckerberg Biohub Investigator. **Competing interests:** The authors declare no conflicts of interest.

10.1126/science.aaw9255

RESEARCH ARTICLE SUMMARY

IMMUNOLOGY

MAIT cells are imprinted by the microbiota in early life and promote tissue repair

Michael G. Constantinides, Verena M. Link, Samira Tamoutounour, Andrea C. Wong, P. Juliana Perez-Chaparro, Seong-Ji Han, Y. Erin Chen, Kelin Li, Sepideh Farhat, Antonin Weckel, Siddharth R. Krishnamurthy, Ivan Vujkovic-Cvijin, Jonathan L. Linehan, Nicolas Bouladoux, E. Dean Merrill, Sobhan Roy, Daniel J. Cua, Erin J. Adams, Avinash Bhandoola, Tiffany C. Scharschmidt, Jeffrey Aubé, Michael A. Fischbach, Yasmine Belkaid*

INTRODUCTION: The microbiota promotes the maturation and homeostasis of the immune system, in part through the release of microbial products. Early-life microbial colonization has been shown to play a fundamental role in the development of the immune system and imparts long-lasting effects on host fitness. However, despite the importance of this early host-microbiota dialogue, little is known about the microbial-derived signals and antigens involved. This question is of particular importance for mucosal-associated invariant T (MAIT) cells, which are predominantly located in tissues colonized by the microbiota and characterized by their recognition of microbial-derived intermediates of vitamin B2 (riboflavin) synthesis. Because riboflavin synthesis is broadly conserved among bacteria and fungi, MAIT cells are thought to be particularly dependent on the microbiota.

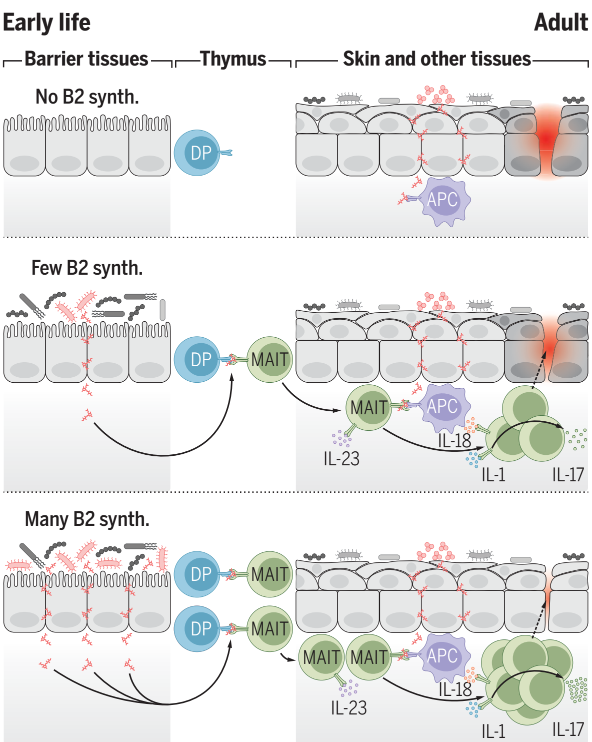
RATIONALE: Although MAIT cells are the predominant innate-like lymphocyte subset in humans, there is remarkable variability in their abundance between individuals. How commensal-derived antigens contribute to the variability in MAIT cell abundance and their function has not been established. Furthermore, the extent to which MAIT cells promote tissue physiology remains to be determined.

RESULTS: MAIT cells were highly abundant within human and mouse skin. However, genetically identical mice housed in distinct cages showed striking variability in the proportion of MAIT cells. By contrast, animals in the same cage had similar frequencies of MAIT cells, supporting the hypothesis that these differences were associated with distinct microbiota. MAIT cells accumulated in barrier tissues between 2 and 3 weeks of age, bolstering the idea that MAIT cells develop during a very specific temporal window and in response to defined microbial exposure. The isolation of early-life intestinal commensals and subsequent colonization of neonatal germ-free mice with defined bacteria induced MAIT cell development. Conversely, colonization later in life failed to promote their development within tissues, indicating that microbial exposure must occur during an early-life window, imprinting MAIT cell abundance for life. Commensals that induced MAIT cell development were capable of synthesizing riboflavin, demonstrating antigen necessity, whereas the development of MAIT cells after treatment with a riboflavin derivative indicated sufficiency. After their development in response to early-life commensals, MAIT cells represented a dominant type-17 effector subset in the skin, and cutaneous MAIT cells distinctly expressed a transcriptional program associated with tissue repair. Cutaneous MAIT cells were tissue-resident and required the cytokine interleukin-23 (IL-23) for their homeostasis. These lymphocytes were capable of responding locally to skin commensals in a manner that required IL-1 and IL-18 as well as antigen presentation mediated by the major histocompatibility complex molecule MR1. MR1-mediated presentation of riboflavin metabolites was necessary and sufficient for MAIT cell recognition of skin commensals and further enhanced the tissue-repair program of these lymphocytes. Within the skin, MAIT cells were distinctly localized at the interface of the dermis and epidermis, in close proximity to the basal layer. Topical application of a riboflavin derivative selectively increased MAIT cells in the skin and was sufficient to promote cutaneous wound healing, demonstrating that MAIT cells contribute to skin physiology.

CONCLUSION: Our work demonstrates how early microbial encounters have long-term effects on the composition of the immune system. We show that MAIT cells are induced during a specific early-life window in response to riboflavin-synthesizing commensals. This phenomenon permanently imprints the abundance of this subset in tissues, controlling tissue repair and homeostasis. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aax6624



Early-life microbial exposure imprints the abundance of MAIT cells, and subsequent interactions with the microbiota modulate their function. (Top) A lack of microbial exposure in early life, results in an absence of MAIT cells in adults. (Middle) Early colonization with a low abundance of vitamin B2–synthesizing (synth.) commensals induces minimal MAIT cell development, whereas (Bottom) a high abundance of these microbes imprints a high frequency of tissue-resident MAIT cells in adults. MAIT cells are selected on double-positive (DP) thymocytes and require IL-23 for their development and/or accumulation. Subsequent recognition of commensals requires MR1-mediated presentation of riboflavin derivatives and IL-18. Microbial stimulation of MAIT cells prompts IL-1–dependent IL-17A production and reinforces a tissue repair program inherent to MAIT cells.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: ybelkaid@niaid.nih.gov
Cite this article as M. Constantinides *et al.*, *Science* **366**, eaax6624 (2019). DOI: 10.1126/science.aax6624

RESEARCH ARTICLE

IMMUNOLOGY

MAIT cells are imprinted by the microbiota in early life and promote tissue repair

Michael G. Constantinides[1], Verena M. Link[1], Samira Tamoutounour[1]*, Andrea C. Wong[2], P. Juliana Perez-Chaparro[3], Seong-Ji Han[1], Y. Erin Chen[4], Kelin Li[5], Sepideh Farhat[6], Antonin Weckel[6], Siddharth R. Krishnamurthy[1], Ivan Vujkovic-Cvijin[1], Jonathan L. Linehan[1]†, Nicolas Bouladoux[1,3], E. Dean Merrill[1], Sobhan Roy[7], Daniel J. Cua[8]‡, Erin J. Adams[7], Avinash Bhandoola[9], Tiffany C. Scharschmidt[6], Jeffrey Aubé[5], Michael A. Fischbach[4], Yasmine Belkaid[1,3]§

How early-life colonization and subsequent exposure to the microbiota affect long-term tissue immunity remains poorly understood. Here, we show that the development of mucosal-associated invariant T (MAIT) cells relies on a specific temporal window, after which MAIT cell development is permanently impaired. This imprinting depends on early-life exposure to defined microbes that synthesize riboflavin-derived antigens. In adults, cutaneous MAIT cells are a dominant population of interleukin-17A (IL-17A)–producing lymphocytes, which display a distinct transcriptional signature and can subsequently respond to skin commensals in an IL-1–, IL-18–, and antigen-dependent manner. Consequently, local activation of cutaneous MAIT cells promotes wound healing. Together, our work uncovers a privileged interaction between defined members of the microbiota and MAIT cells, which sequentially controls both tissue-imprinting and subsequent responses to injury.

Every barrier site harbors a distinct community of commensal microbes, known as the microbiota, that controls host physiology (*1*). These microbes promote the maturation and homeostasis of the immune system, in part through the release of microbial products, including metabolites such as short-chain fatty acids, aryl hydrocarbon receptor (AhR) ligands, and polyamines (*2*). In turn, the immune system modulates the composition of the microbiota and maintains the segregation of commensals by sustaining barrier tissue function (*1*).

Early-life microbial colonization has been shown to play a fundamental role in the development of the immune system and imparts long-lasting effects on host fitness. For example, neonatal colonization of the lungs promotes tolerance to allergens in adult mice through the induction of regulatory T cells (T_{reg} cells) (*3*), whereas colonization of neonatal skin similarly induces T_{reg} cell–mediated tolerance to commensal microbes (*4*). During weaning, the microbiota induces a transient up-regulation of phosphorylated signal transducer and activator of transcription 3 (STAT3) in epithelial cells and innate lymphoid cells (ILCs) (*5*), as well as an increase in colonic T_{reg} cells (*6*), which can affect host metabolism and susceptibility to inflammatory disorders. High microbial diversity within the intestine during early life is necessary to impede immunoglobulin E (IgE) class switching of mucosal B cells, reducing susceptibility to orally induced anaphylaxis (*7*). Furthermore, early colonization with the intestinal commensal *Bacteroides fragilis* inhibits the proliferation of natural killer T (NKT) cells, which prevents oxazolone-induced colitis in adult mice (*8*, *9*). However, despite the importance of this early host-microbiota dialogue, little is known about the microbially derived signals and antigens involved. This question is of particular importance for unconventional T cells, which are predominantly located in tissues colonized by the microbiota and characterized by their recognition of conserved antigenic motifs.

Exhibiting characteristics of both innate and adaptive immunity, innate-like lymphocytes such as mucosal-associated invariant T (MAIT) cells, NKT cells, and γδ T cells acquire their effector functions during development, which direct their localization to tissues and enable them to respond immediately upon primary antigen recognition (*10*). Consequently, within barrier sites, innate-like lymphocytes are likely poised to respond to the microbiota and play a dominant role in mediating host-commensal interactions. Both γδ T cells and $CD8^+$ T cells restricted by the nonclassical major histocompatibility complex (MHC)–Ib molecule H2-M3 respond to skin commensals (*11*, *12*), whereas γδ T cells can mediate both beneficial and deleterious effects of the microbiota in the intestines and lungs (*13*, *14*). MAIT cells express semi-invariant αβ T cell receptors that recognize microbial-derived intermediates of vitamin B2 (riboflavin) synthesis presented by the MHC-Ib molecule MR1 (*15*). Because riboflavin synthesis is broadly conserved among bacteria and fungi (*16*), MAIT cells are thought to be particularly dependent on the microbiota. Germ-free (GF) mice exhibit fewer MAIT cells than do animals housed in specific-pathogen-free (SPF) conditions (*17*, *18*). Although these cells are the predominant innate-like lymphocyte subset in humans, where they compose up to 45% of hepatic lymphocytes (*19*), surprisingly little is known about them. Although the function of MAIT cells remains largely unclear, in humans these cells can be altered in defined inflammatory or infectious settings (*20*), and their frequencies show remarkable variability among individuals (*21*, *22*). Although the paucity of MAIT cells in mice has rendered the exploration of their function challenging, MAIT cells have been proposed to promote inflammation and microbial defense (*23*–*25*). How commensal-derived antigens contribute to the abundance and function of MAIT cells has not been established. Furthermore, the extent to which MAIT cells promote tissue physiology remains to be determined.

Here, we show that the development of MAIT cells and their long-term frequencies within tissues depend on a very specific developmental window and, in particular, early-life exposure to defined microbial communities enriched in riboflavin-synthesizing bacteria, such as Enterobacteriaceae. This defined microbial exposure must occur during the first few weeks of life, after which MAIT cell development is permanently impaired. After their development, MAIT cells become a dominant, tissue-resident population of interleukin-17A (IL-17A)–producing lymphocytes within the skin and display a transcriptional profile that is distinct from MAIT cells in other tissues. Cutaneous MAIT cells can subsequently respond to defined skin commensals or commensal-derived metabolites in an antigen-dependent manner. Consequently, the local activation of cutaneous MAIT cells promotes wound healing. This work therefore uncovers a privileged interaction between defined members of the microbiota and MAIT cells, resulting in a sequential

[1]Metaorganism Immunity Section, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [2]Immunology Graduate Group, University of Pennsylvania, Philadelphia, PA 19104, USA. [3]NIAID Microbiome Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [4]Department of Bioengineering and ChEM-H, Stanford University, Stanford, CA 94305, USA. [5]Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA. [6]Department of Dermatology, University of California, San Francisco, CA 94143, USA. [7]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA. [8]Merck & Co., Merck Research Laboratories, Palo Alto, CA 94304, USA. [9]Laboratory of Genome Integrity, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
*Present address: L'Oréal Research and Innovation, Aulnay-sous-Bois, 93600, France. †Present address: Department of Cancer Immunology, Genentech, South San Francisco, CA 94080, USA. ‡Present address: Janssen Research and Development, Spring House, PA 19477, USA.
§Corresponding author. Email: ybelkaid@niaid.nih.gov

dialogue that controls both tissue-imprinting and subsequent responses to injury.

Early-life exposure to riboflavin-synthesizing commensals is required for MAIT cell development

Although MAIT cells are abundant in human tissues (19), they are comparatively infrequent in mouse organs, where they typically represent less than 1% of αβ T cells (26). Because the skin is characterized by an unusually high frequency of unconventional T cells (Fig. 1A), we assessed the frequency of MAIT cells at this barrier site using an MR1 tetramer loaded with the riboflavin derivative 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU) (27). Although nonspecific MR1 tetramer staining was minimal, as demonstrated by use of a MR1 tetramer loaded with 6-formylpterin (6-FP) (fig. S1A), we used 5-OP-RU–loaded MR1 tetramers conjugated to two different fluorochromes for increased specificity, when possible (fig. S1B). This analysis revealed that MAIT cells were highly enriched in the skin (Fig. 1, B and C), where they composed up to 40% of αβ T cells (Fig. 1, D and E). The frequency of MAIT cells in the skin correlated highly with their frequency in other organs (fig. S1C). Furthermore, MAIT cells were also enriched in human skin compared with the reported ~1% in peripheral blood (20), averaging ~2% of $CD3^+$ lymphocytes (Fig. 1F). Thus, MAIT cells represent a substantial population of skin lymphocytes, both in mice and humans.

A comparison of mice housed in conventional SPF conditions to GF animals confirmed previous observations that the microbiota is necessary for the presence of MAIT cells in the thymus, spleen, and gut (17, 18) and extended this requirement for MAIT cells in the skin, lungs, and liver (Fig. 1B). Although MAIT cells were highly abundant within the skin, wild-type (WT) SPF mice housed in distinct cages showed striking variability in the proportion of MAIT cells among genetically identical animals, ranging from <5 to 40% of αβ T cells (Fig. 1, D and E). By contrast, animals in the same cage had similar frequencies of MAIT cells in the skin and other organs (Fig. 1E), supporting the hypothesis that these differences were associated with distinct microbiota.

To establish how the microbiota could modulate MAIT cell development and/or accumulation, we first assessed the kinetics of unconventional T cell accumulation within barrier tissues. Although γδ T cells and NKT cells were present in barrier tissues of 2-week-old mice, there was a dearth of MAIT cells in the skin, lungs, and intestines (Fig. 1G). MAIT cells accumulated in all barrier tissues assessed by 3 weeks of age, supporting the idea that MAIT cell development occurs during a very specific developmental window and in response to defined microbial exposure. Therefore, we longitudinally assessed the intestinal microbiota of WT mice during early life using 16*S* ribosomal RNA (rRNA) gene sequencing (Fig. 1G). At 2 weeks of age, the intestinal microbiota was predominantly composed of Lactobacillaceae, with a transient enrichment of Enterobacteriaceae (Fig. 1G), which is consistent with previous work that has identified an abundance of these two bacterial families in early life (28). Isolation of several bacterial strains from 2-week-old mice, which developed MAIT cells as adults (fig. S1, D and E), allowed us to generate an early-life microbial community composed of two Lactobacillaceae members (*Lactobacillus murinus* and *Lactobacillus johnsonii*), two Enterobacteriaceae members (*Proteus mirabilis* and *Klebsiella oxytoca*), and the common intestinal commensal *Enterococcus faecalis*. To determine whether this five-species community was sufficient to promote MAIT cell development, we colonized GF mice by administering the bacteria at weeks 1, 2, and 3 by means of oral gavage (Fig. 1H). The colonization of neonatal GF mice with our early-life microbial community induced MAIT cell development to the same extent observed in SPF mice harboring high MAIT cell frequencies (Fig. 1I). Thus, this limited community was sufficient to provide the necessary developmental and/or survival signals. Conversely, exposure later in life failed to promote MAIT cell development and accumulation within tissues (Fig. 1, H and I), while still inducing T helper 17 (T_H17) and IL-17A$^+$ γδ T cells (fig. S1F). Similarly, conventionalized mice (adult GF mice cohoused with SPF animals) remained deficient in MAIT cells (Fig. 1I). Thus, exposure to the microbiota during an early-life window has a profound impact on MAIT cell development, and exposure later in life cannot compensate for a lack of neonatal colonization.

Whole-genome sequencing of the five-species community indicated that *P. mirabilis* and *K. oxytoca* have the genes *ribA*, *ribD*, and *pyrp2*, which encode the riboflavin synthesis enzymes necessary to generate the MAIT cell antigen 5-OP-RU, whereas the *Lactobacillus* species and *E. faecalis* do not (Fig. 1J), suggesting that, within our defined community, early colonization with Enterobacteriaceae were responsible for the emergence of the MAIT cell lineage. The monocolonization of neonatal GF mice with *P. mirabilis*, but not *L. johnsonii*, was sufficient to drive MAIT cell development (Fig. 1K). Segmented filamentous bacteria (SFB), a known inducer of T_H17 cells that lacks riboflavin synthesis enzymes (Fig. 1J and fig. S1G) (29, 30), also failed to induce MAIT cells (Fig. 1K). As we observed with the five-species early-life community, the ability of *P. mirabilis* to promote MAIT cell development was dependent on neonatal colonization or exposure from birth, both in the skin and other tissues (fig. S1H). *P. mirabilis* colonization of both GF neonates and adults was sufficient to induce mature $CD24^-$ $CD44^+$ MAIT cells in the thymus (fig. S1I), indicating that microbial exposure in early life is required for the development of tissue MAIT cells but not the thymic developmental stages. This corroborates recent work that demonstrates that microbially derived 5-OP-RU can promote the thymic development of MAIT cells in GF mice (31). Enterobacteriaceae, including both *Proteus* and *Klebsiella* species, are enriched in the gut microbiota of both human and mouse neonates and decrease over time (28, 32, 33). Thus, the early abundance of Enterobacteriaceae may play an essential role in the long-term imprinting of MAIT cells.

We next assessed whether exposure to microbial antigens alone during early life was sufficient to induce MAIT cell development and whether such exposure could occur at other barrier sites, such as the skin. To this end, we administered a mixture of 5-amino-6-D-ribitylaminouracil (5-A-RU) and methylglyoxal, which react to form the unstable MAIT cell ligand 5-OP-RU (34), to the skin of GF mice on a weekly basis, beginning at 1 week of age. Topical application of 5-OP-RU was sufficient to induce a significant increase in the number of MAIT cells within the skin (Fig. 1L), without promoting the accumulation of other T cell subsets (fig. S1J). This demonstrated that antigen recognition in the absence of microbial-induced proinflammatory cytokines was adequate for MAIT cell development. In agreement with our observation that adult microbial colonization was insufficient to promote MAIT cell development (Fig. 1I), topical application of 5-OP-RU to the skin of adult GF mice failed to induce MAIT cells (Fig. 1L). Thus, early-life exposure to commensals that produce defined metabolites, or to microbially derived metabolites alone, is required for MAIT cell development and their presence later in life. Although this exposure to microbial antigens can occur at multiple barrier sites, it must take place during a defined period in early life. If not, MAIT cell development in tissues is permanently impaired.

Cutaneous MAIT cells express a type-17 transcriptional program and require homeostatic IL-23

MAIT cells have previously been shown to produce either type-1– or type-17–associated cytokines (18, 26). After stimulation, skin MAIT cells robustly produced IL-17A but not interferon-γ (IFNγ) (Fig. 2A), suggesting that within the skin, these cells are exclusively type-17 lymphocytes. As such, in mice with high MAIT cell frequencies, MAIT cells accounted for the largest proportion of IL-17A$^+$ αβ T cells (Fig. 2B), indicating that these innate-like lymphocytes can represent a dominant type-17 effector subset in the skin.

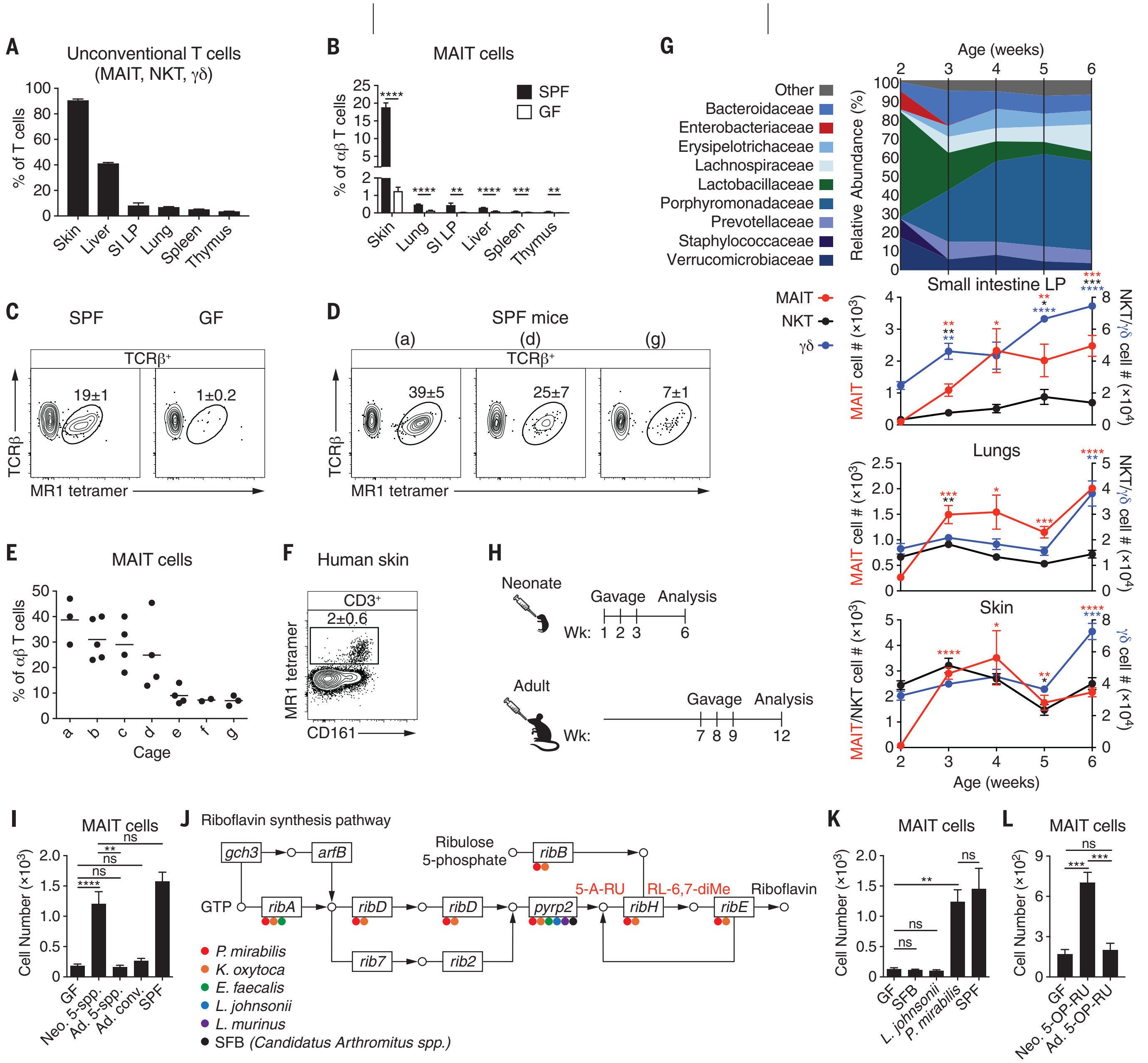


Fig. 1. Early-life exposure to riboflavin-synthesizing commensals is required for MAIT cell development. (**A**) Percentage of T cells (TCRβ$^+$ or TCRγδ$^+$) that are MAIT, NKT, or γδ T cells in WT mice. SI LP, small intestine lamina propria. (**B**) Percentage of MAIT cells among αβ T cells in SPF and GF WT mice. (**C**) Flow cytometry of TCRβ$^+$ lymphocytes from the skin of SPF and GF WT mice. (**D** and **E**) Analysis of SPF mice housed in different cages (denoted a to g). (D) Flow cytometry of TCRβ$^+$ lymphocytes from the skin of mice housed in the indicated cages. (E) The percentage of MAIT cells among αβ T cells, with each dot representing an individual mouse. (**F**) Flow cytometry of CD3$^+$ lymphocytes from a human skin biopsy. (**G**) 16*S* rRNA gene sequencing of feces from WT SPF mice longitudinally sampled from 2 to 6 weeks of age and the number of MAIT, NKT, and γδ T cells present in WT SPF mice at the corresponding ages. Asterisks denote statistically significant changes in cell number from 2 weeks of age. (**H**) Neonatal GF mice received oral gavages at 1, 2, and 3 weeks after birth, whereas adult GF mice received oral gavages at 7, 8, and 9 weeks of age. Both were analyzed 5 weeks after the initial gavage. (**I**) Number of MAIT cells in the skin of GF mice administered oral gavages of the five-species (5-spp.) community (*P. mirabilis*, *K. oxytoca*, *E. faecalis*, *L. johnsonii*, and *L. murinus*) either as neonates (Neo.) or adults (Ad.) as depicted in (H) or conventionalized (conv.) by cohousing with SPF mice for 3 weeks as adults. (**J**) Presence of riboflavin synthesis genes in bacteria of the five-species community and SFB denoted with the appropriate color. 5-A-RU, which reacts with methylglyoxal to form the MAIT cell antigen 5-OP-RU, and 6,7-dimethyl-8-(1-D-ribityl)lumazine (RL-6,7-diMe) are denoted in red. (**K**) Number of MAIT cells in the skin of GF mice monocolonized as neonates, with the indicated bacterial species compared with GF and SPF controls. (**L**) PBS solution of 1 mM 5-A-RU and 25 mM methylglyoxal (referred to as 5-OP-RU) was topically applied once per week to the skin of GF mice beginning either at 1 week of age (Neo.) or 7 weeks of age (Ad.). Number of MAIT cells in the skin was assessed 5 weeks after the initial application. Flow cytometry gate frequencies and graphs indicate means ± SEM. Data represent at least two experiments with four or more mice per group. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$ as calculated with Student's *t* test. "ns" denotes that comparison was not significant.

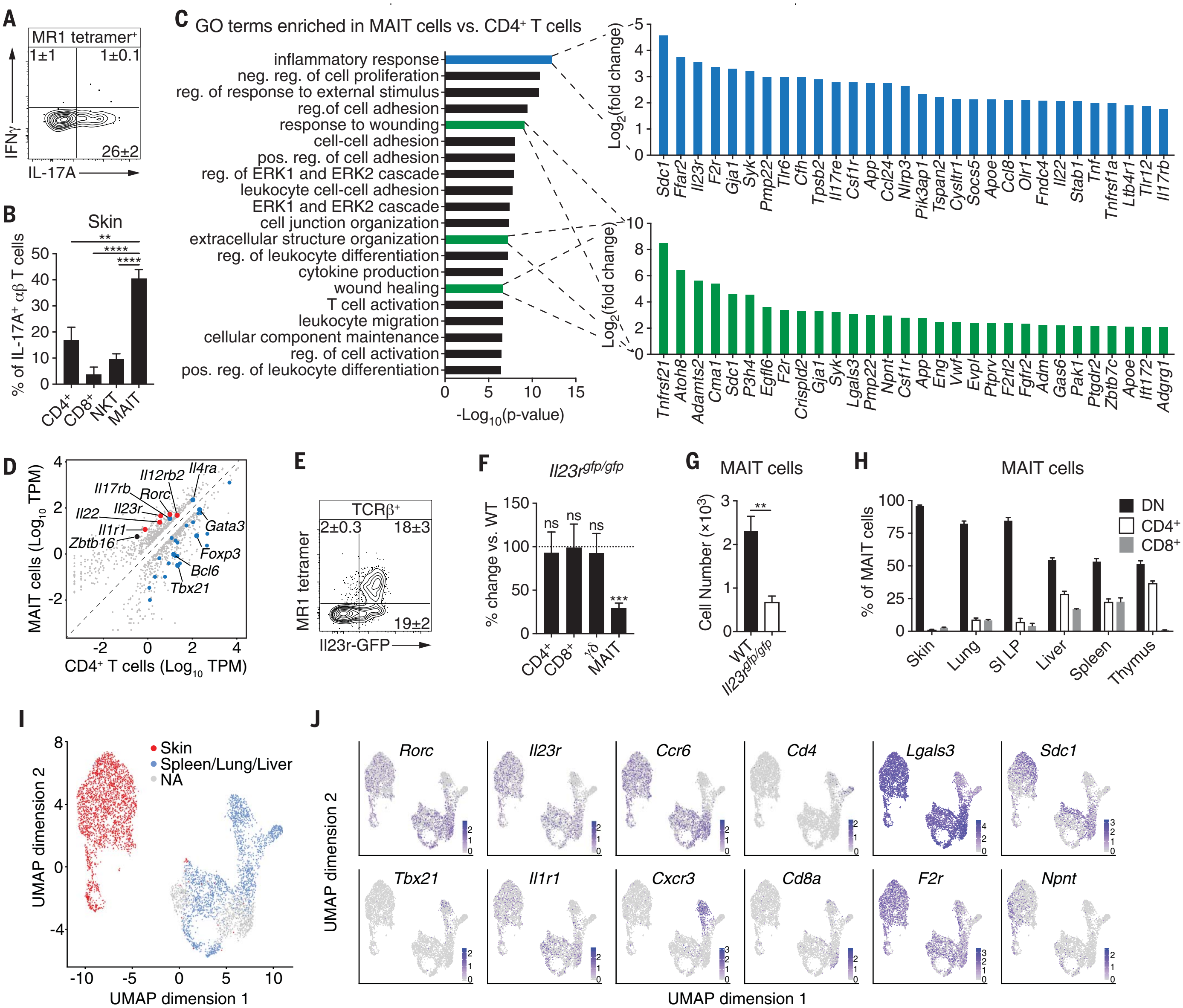


Fig. 2. Cutaneous MAIT cells express a type-17 transcriptional program and require homeostatic IL-23. (**A**) Flow cytometry of cytokine production by MAIT cells from the skin of WT mice. (**B**) Percentage of IL-17A$^+$ αβ T cells within WT mouse skin that were CD4$^+$, CD8$^+$, NKT, or MAIT cells. (**C** and **D**) RNA-seq data of cutaneous MAIT cells and CD4$^+$ CD25$^-$ T cells. (C) Top 20 GO terms that were enriched in MAIT cells and, for the indicated GO terms, the top 30 genes up-regulated in MAIT cells. Positive (pos.), negative (neg.), and regulation (reg.) are abbreviated. (D) Expression plot normalized to transcripts per million (TPM), with a minimum fold-change of 2 and FDR < 5%. Type-17 genes are denoted in red, and genes associated with Type-1, Type-2, T regulatory, and T helper programs are highlighted in blue. (**E**) Representative flow cytometry plot of TCRβ$^+$ lymphocytes from the skin of *Il23r*$^{gfp/+}$ mice. (**F**) Percentage change of the indicated T cells in the skin of *Il23r*$^{gfp/gfp}$ mice compared with WT controls. Statistics denote whether the percentage differs significantly from the WT mean (100%). (**G**) Number of cutaneous MAIT cells in *Il23r*$^{gfp/gfp}$ mice and WT controls. (**H**) Percentage of MAIT cells in the indicated organs that expressed CD4, CD8, or neither coreceptor [double negative (DN)]. (**I** and **J**) scRNA-seq data of MAIT cells sorted from mouse skin, spleen, lung, and liver. Clusters were assigned to "Skin" or "Spleen/Lung/Liver" on the basis of the presence of hashtag oligonucleotides (HTOs) from those tissues. Clusters that did not have a predominance of HTOs from either skin or the other tissues were not assigned ("NA"). (I) UMAP plot displaying the distribution of MAIT cell clusters assigned to skin and the other tissues. (J) UMAP plots depicting expression of the indicated transcripts. Flow cytometry gate frequencies and graphs indicate means ± SEM. Data represent at least two experiments with four or more mice per group. $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$ as calculated with Student's *t* test. "ns" denotes that comparison was not significant.

To establish the gene expression profile of cutaneous MAIT cells, we compared them with conventional CD4$^+$ T cells from the skin using RNA-sequencing (RNA-seq) (fig. S2A). Gene Ontology (GO) enrichment analysis revealed that cutaneous MAIT cells express a transcriptional profile consistent with T cell activation, including leukocyte differentiation, migration and cell adhesion, cytokine production, and the extracellular signal–regulated kinase–1 and –2 (ERK1/2) cascade, which is downstream of T cell receptor (TCR) signaling (Fig. 2C). MAIT cells were also enriched in GO terms associated with tissue repair (Fig. 2C), suggesting that these lymphocytes may contribute

to the maintenance or restoration of tissue homeostasis.

As expected, cutaneous MAIT cells expressed the gene *Zbtb16* (Fig. 2D), which encodes the transcription factor promyelocytic leukemia zinc finger (PLZF) and is necessary for their development (*18*, *26*). Consistent with their IL-17A potential, MAIT cells expressed type-17–associated transcripts, including *Il22* and *Rorc*, the gene for the transcription factor retinoic acid receptor–related orphan receptor–γt (RORγt), as well as *Il1r1* and *Il23r*, which encode receptors for IL-1 and IL-23, respectively (Fig. 2D). Analysis of $Il1r1^{-/-}$ mice indicated that cutaneous MAIT cells do not require IL-1 signaling for their development (fig. S2, B and C). MAIT cells expressed the highest levels of IL-23R among αβ T cells (Fig. 2E), and IL-23 receptor–deficient mice ($Il23r^{gfp/gfp}$) exhibited fewer MAIT cells than did WT animals (Fig. 2, F and G, and fig. S2, D and E), supporting the idea that homeostatic IL-23 contributes to MAIT cell development and/or accumulation within tissues. Although SFB is known to promote IL-23 expression (*35*), SFB did not induce MAIT cells in GF mice (Fig. 1K), indicating that IL-23 is insufficient to promote MAIT cell development in the absence of microbial-derived metabolites. Thus, MAIT cells constitute a substantial IL-17A–producing population within the skin that requires IL-23 signaling for development and/or survival.

Although the MAIT cell TCR does not require the CD4 or CD8 coreceptors to facilitate recognition of antigens presented by MR1, $CD4^+$, $CD8^+$, and $CD4^-$ $CD8^-$ double-negative (DN) MAIT cells have been described in mice (*26*), whereas $CD8^+$ and DN subsets exist in humans (*36*). We found that MAIT cells from the skin and other barrier sites, such as the lungs and gut, were almost exclusively DN (Fig. 2H). Conversely, a large proportion of MAIT cells in the thymus, spleen, and liver expressed either CD4 or CD8. In humans, DN MAIT cells produce more IL-17 and can arise from $CD8^+$ MAIT cells in vitro (*36*), suggesting that the mouse DN MAIT cells at barrier tissues may represent the most mature developmental stage.

We next compared the transcriptional profile of MAIT cells from the skin to other tissues by means of single-cell RNA-sequencing (scRNA-seq). Cutaneous MAIT cells displayed a markedly different transcriptional profile (Fig. 2, I and J, and fig. S2, F and G). In concordance with our previous data (Fig. 2, D and H), skin MAIT cells expressed *Rorc*, *Il1r1*, and *Il23r* and predominantly lacked expression of *Cd4* and *Cd8a* (Fig. 2J). Type-1 MAIT cells expressing *Tbx21*, which encodes T-bet, and *Cxcr3* were present in nonskin tissues, whereas the $Rorc^+$ MAIT cells expressed *Ccr6* (Fig. 2J). Additionally, scRNA-seq confirmed that cutaneous MAIT cells express multiple genes associated with tissue repair, including *Lgals3*, *F2r*, *Sdc1*, and *Npnt* (Fig. 2J). MAIT cells expressing either coreceptor did not cluster distinctly from DN MAIT cells from the same tissue (Fig. 2J), suggesting that the transcriptional differences between the $CD4^+$, $CD8^+$, and DN subsets did not outweigh their tissue-specific programs.

Skin-resident MAIT cells respond to cutaneous microbes in an IL-1– and IL-18–dependent manner

The skin is readily exposed to the environment and colonized by a diverse microbial community that is second in size only to the intestinal microbiome (*37*). We have previously shown that topical application of the skin commensal *Staphylococcus epidermidis* promotes the accumulation of conventional $CD4^+$ T cells, H2-M3–restricted $CD8^+$ T cells, and γδ T cells within the skin without causing inflammation (*11*, *12*, *38*). Although the relative abundance of MAIT cells within tissues is imprinted through early exposure to defined microbes, MAIT cells may also subsequently respond to alterations in the abundance or composition of tissue-resident microbiota—specifically, those capable of synthesizing riboflavin. Using strains of *S. epidermidis* that either induce nonclassical $CD8^+$ T cells or not (fig. S3A) (*12*), we observed that in both cases, topical association expanded the cutaneous DN MAIT cell population and increased their production of IL-17A (Fig. 3, A to D, and fig. S3, B to D). Furthermore, there was a strong positive correlation between the abundance of MAIT cells before and after association with *S. epidermidis* (fig. S3, E and F), indicating that SPF animals that were developmentally imprinted with a low MAIT cell frequency will retain a lower level of MAIT cells after subsequent interactions with the microbiota.

To determine whether the MAIT cell response to *S. epidermidis* resulted from local expansion or required priming in the lymph node, we associated $Lta^{-/-}$ mice (*39*). Although the number of $CD4^+$, $CD8^+$, and γδ T cells after *S. epidermidis* association were significantly decreased in the absence of lymph nodes as expected (Fig. 3E), *S. epidermidis* still promoted an expansion of MAIT cells in the skin of LTα-deficient mice. This suggested that MAIT cells respond to skin commensals locally, either through antigen recognition and/or cytokine stimulation. To further interrogate whether MAIT cells remain within the skin tissue after stimulation, we associated congenically labeled mice with *S. epidermidis* and performed parabiosis (Fig. 3F). As previously described for NKT cells, subsets of γδ T cells, and MAIT cells in other tissues (*40*–*43*), unconventional T cells in the skin, including MAIT cells, were host-derived (Fig. 3G). Thus, MAIT cells persist in the skin as a tissue-resident population and respond locally to skin commensals.

MAIT cells have been shown to respond to cytokine stimulation in an antigen-independent manner (*44*). More generally, the relative contribution of antigen versus cytokines in MAIT cell activation in vivo remains poorly understood. Although homeostatic IL-23 contributed to early MAIT cell development (Fig. 2, F and G, and fig. S2E), blocking the IL-23 receptor during *S. epidermidis* association did not reduce MAIT cell numbers (Fig. 3H). Because MAIT cells are known to up-regulate IL-18Rα during development (*18*), we confirmed that cutaneous MAIT cells expressed IL-18Rα (Fig. 3I). Neutralization of IL-18 significantly reduced the MAIT cell response to *S. epidermidis* association (Fig. 3, J to L, and fig. S3G), indicating that IL-18 signaling was required for optimal MAIT cell expansion. We have previously shown that topical application of *S. epidermidis* elicits the release of IL-1α, which promotes IL-17 responses by γδ and conventional αβ T cells (*45*). Although IL-1 receptor–deficient animals ($Il1r1^{-/-}$) had equivalent numbers of cutaneous MAIT cells as WT mice at steady state and upon *S. epidermidis* application (Fig. 3M), their ability to produce IL-17A was significantly decreased (Fig. 3, N and O). Thus, IL-18 was necessary for the expansion of MAIT cells in response to *S. epidermidis*, and local IL-1 signaling was required for the licensing of IL-17A production.

We next assessed how *S. epidermidis* association affects the transcriptional profile of tissue-resident MAIT cells. RNA-seq analysis revealed that application of *S. epidermidis* led to an enrichment in GO terms associated with leukocyte activation and tissue repair (Fig. 3, P and Q), indicating that skin commensals may reinforce the functional capabilities of cutaneous MAIT cells. In addition to inducing the expression of genes downstream of TCR signaling, such as *Egr1* and *Egr2* (*46*), *S. epidermidis* association promoted the expression of genes associated specifically with angiogenesis, including angiogenin (*Ang*) and *Hgf* (*47*, *48*), as well as genes that more broadly promote tissue repair, such as *Ceacam1*, *Grn*, *Hmox1*, *Igf1*, and *Pdgfa* (Fig. 3R) (*49*–*53*). Thus, metabolites produced by defined skin commensals may promote the local expansion and activation of cutaneous MAIT cells and enhance their tissue repair program.

MR1-mediated presentation of riboflavin metabolites is necessary and sufficient for MAIT cell recognition of skin commensals

Although the intestine harbors multiple bacterial families that are unable to generate riboflavin, riboflavin synthesis is highly conserved among skin commensals (Fig. 4A). Nearly all Staphylococcaceae species, including our *S. epidermidis* isolates, express riboflavin synthesis genes (Fig. 4, A and B). To assess whether recognition of riboflavin metabolites was necessary for MAIT cell responses within the skin, we generated a mutant strain of *S. epidermidis*

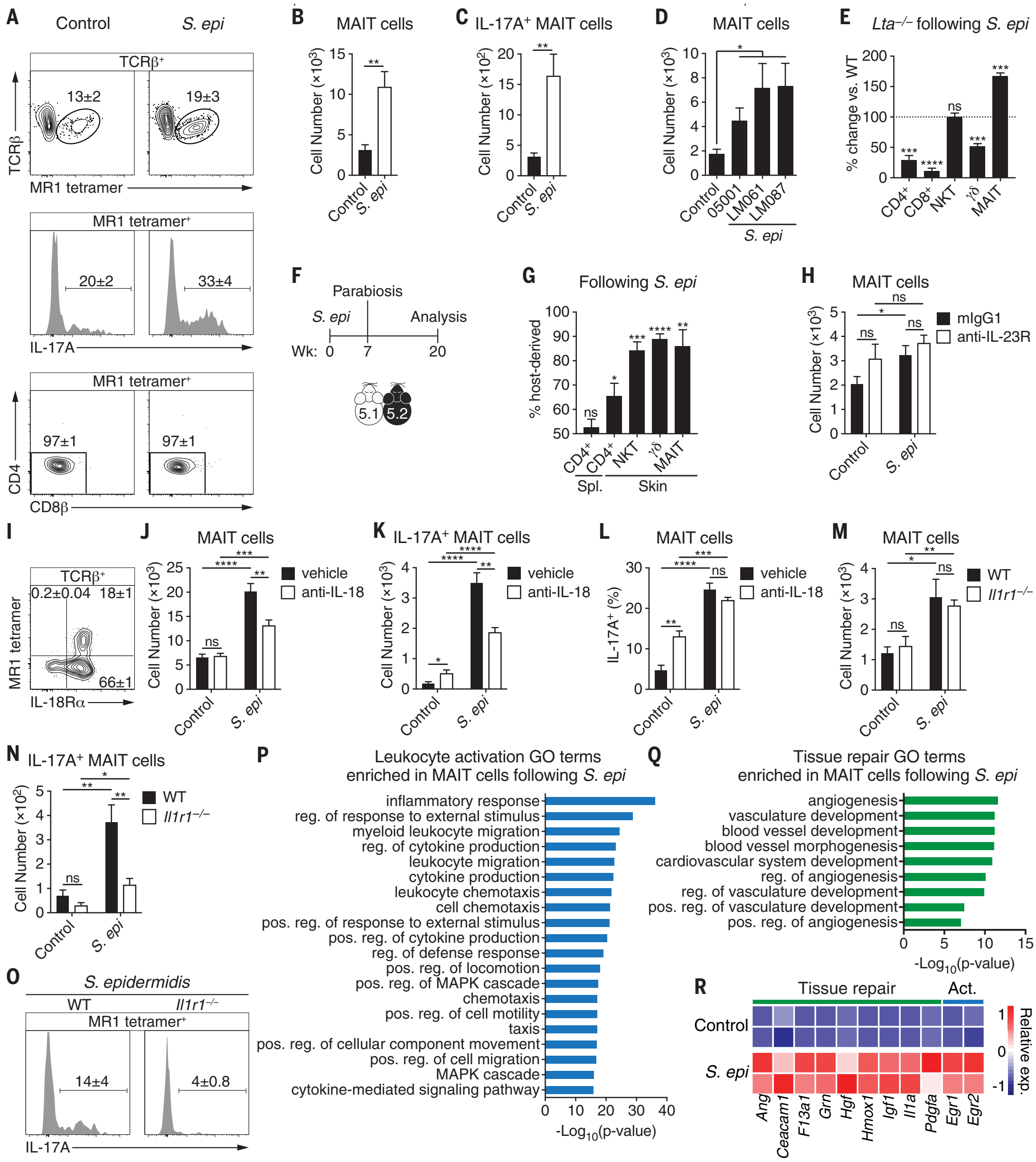


Fig. 3. Skin-resident MAIT cells respond to cutaneous microbes in an IL-1– and IL-18–dependent manner. (**A** to **C**) *S. epidermidis* (*S. epi*) LM061 was topically applied to the skin of WT mice on days 0, 2, 4, and 6. Animals were compared with unassociated (control) mice 14 days after the initial application. (A) Flow cytometry of TCRβ$^+$ lymphocytes (top), IL-17A production by MAIT cells (middle), and coreceptor expression of MAIT cells (bottom) within the skin. (B) Number of cutaneous MAIT cells. (C) Number of cutaneous IL-17A$^+$ MAIT cells. (**D**) Number of cutaneous MAIT cells after topical association with CD8$^+$ T cell–inducing (LM087) and –noninducing (05001 and LM061) strains of *S. epidermidis*. (**E**) Both *Lta*$^{-/-}$ and WT mice were topically associated with *S. epidermidis* LM087, and the percentage change of T cells in the skin of *Lta*$^{-/-}$ mice compared with WT controls is depicted. Statistics denote whether the percentage differs significantly from the WT mean (100%). (**F**) CD45.1 (5.1) and CD45.2 (5.2) mice were topically associated with *S. epidermidis* LM087, conjoined 7 weeks later, and analyzed 13 weeks after parabiosis. (**G**) Frequency of T cells in the indicated tissues of parabiotic mice that are host-derived. (**H**) WT mice were injected subcutaneously with 1 mg of either antibody to IL-23R or mIgG1 isotype control 2 days before the initial application of *S. epidermidis* LM061 on day 0 and again on day 6. Number of cutaneous MAIT cells is depicted. (**I**) Flow cytometry of TCRβ$^+$ lymphocytes from the skin of WT mice. (**J** to **L**) WT mice were injected intraperitoneally with either 1 mg of antibody to IL-18 or saline (vehicle) 2 days before the initial application of *S. epidermidis* LM061 on day 0 and again on days 1, 5, 8, and 11. (J) Number of cutaneous MAIT cells and (K) IL-17A$^+$ MAIT cells and (L) percentage of MAIT cells that are IL-17A$^+$ in mice treated with antibody to IL-18 and control mice (vehicle) that were associated with *S. epidermidis* LM061. (**M** and **N**) Number of cutaneous (M) MAIT cells and (N) IL-17A$^+$ MAIT cells in *Il1r1*$^{-/-}$ mice and WT controls associated with *S. epidermidis* LM061. (**O**) Flow cytometry of IL-17A production by MAIT cells within the skin of *S. epidermidis*–associated mice. (**P** to **R**) RNA-seq data of cutaneous MAIT cells from mice associated with *S. epidermidis* LM061 and unassociated controls. GO terms enriched in MAIT cells from *S. epidermidis*–associated mice that are related to (P) leukocyte activation and (Q) tissue repair. Positive (pos.), negative (neg.), and regulation (reg.) are abbreviated. (R) Heatmap depicting relative expression (exp.) of genes associated with tissue repair and leukocyte activation (Act.). Flow cytometry gate frequencies and graphs indicate means ± SEM. Data represent at least two experiments with four or more mice per group. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$ as calculated with Student's t test. "ns" denotes that comparison was not significant.

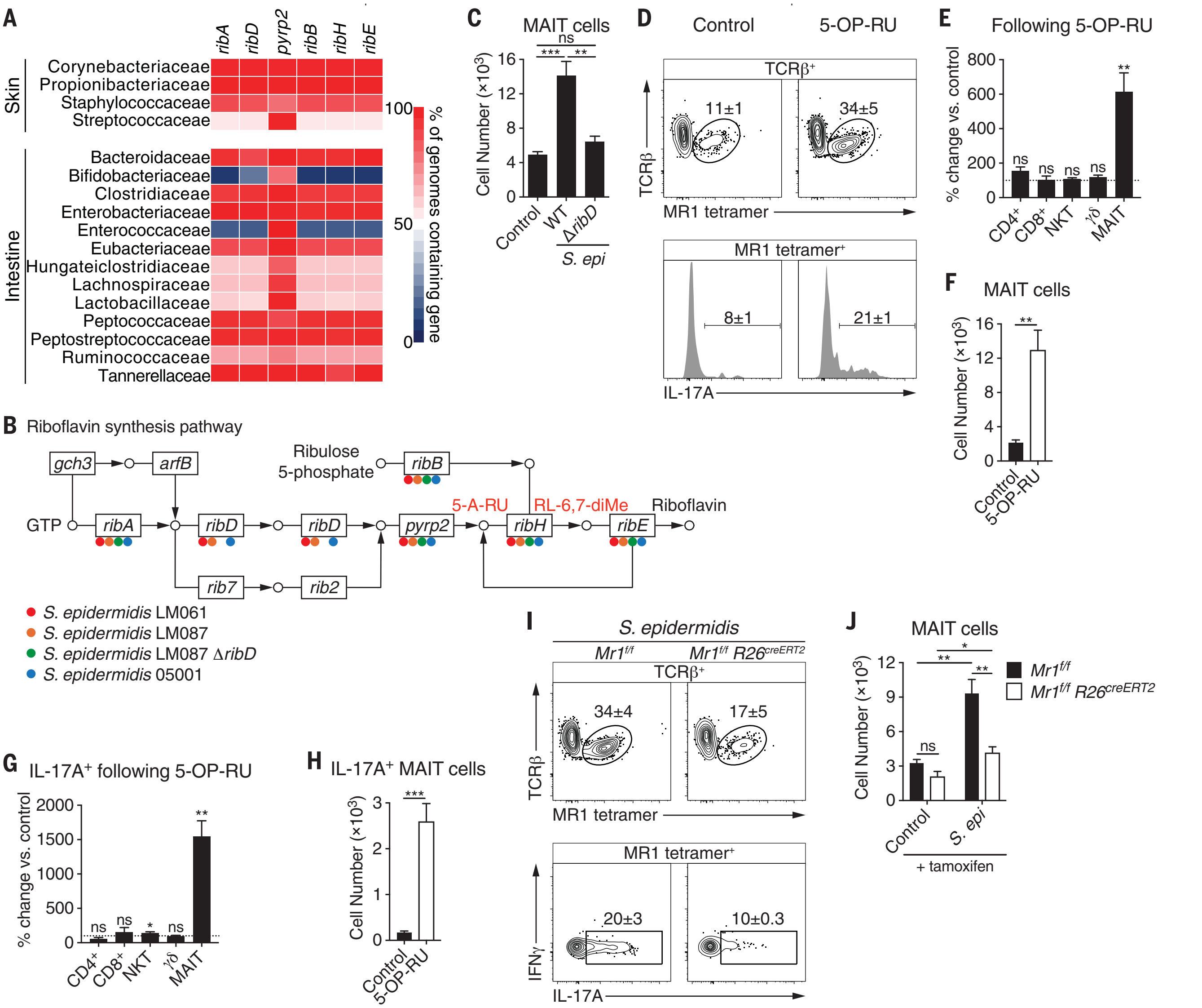


Fig. 4. MR1-mediated presentation of riboflavin metabolites is necessary and sufficient for MAIT cell recognition of skin commensals. (**A**) Heatmap displaying the abundance of riboflavin synthesis genes among species in the indicated bacterial families. (**B**) Presence of riboflavin synthesis genes in strains of *S. epidermidis* denoted with the appropriate color. (**C**) Number of cutaneous MAIT cells in WT mice associated with either *S. epidermidis* LM087 or mutant *S. epidermidis* LM087 Δ*ribD*. (**D** to **H**) 5-OP-RU was topically applied to the skin of WT mice on days 0, 2, 4, and 6. Animals were compared with untreated (control) mice 14 days after the initial application. (D) Flow cytometry of TCRβ$^+$ lymphocytes (top) and IL-17A production by MAIT cells (bottom) within the skin of 5-OP-RU–treated and control mice. Shown in (E) and (G) are the percentage change of the indicated (E) T cells and (G) IL-17A$^+$ T cells in the skin of mice treated with 5-OP-RU compared with untreated controls. Statistics denote whether the percentage differs significantly from the control mean (100%). Shown in (F) and (H) are the number of cutaneous (F) MAIT cells and (H) IL-17A$^+$ MAIT cells in 5-OP-RU–treated and untreated control mice. (**I** and **J**) *Mr1*$^{f/f}$ *R26*^{creERT2} and *Mr1*$^{f/f}$ littermate controls were injected intraperitoneally with 3 mg of tamoxifen 8, 6, 4, and 2 days before the initial association with *S. epidermidis* LM061 and were analyzed 14 days later. (I) Flow cytometry of TCRβ$^+$ lymphocytes and cytokine production by MAIT cells. (J) Number of cutaneous MAIT cells. Flow cytometry gate frequencies and graphs indicate means ± SEM. Data represent at least two experiments with four or more mice per group. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001 as calculated with Student's *t* test. "ns" denotes that comparison was not significant.

with a deletion of the *ribD* gene, which is required for the second step of the riboflavin synthesis pathway, preceding generation of 5-OP-RU (27). Although this mutation prevented riboflavin synthesis from guanosine-5′-triphosphate (GTP), the *S. epidermidis* Δ*ribD* mutant retained the *ribB*, *ribH*, and *ribE* genes necessary to produce MAIT cell antigens from ribulose 5-phosphate (Fig. 4B). WT and Δ*ribD* mutant strains colonized the skin equivalently (fig. S4A). However, in contrast with the WT *S. epidermidis* strain, association with the *S. epidermidis* Δ*ribD* mutant failed to induce a MAIT cell response (Fig. 4C), even though the CD8$^+$ and γδ T cell responses were conserved (fig. S4, B and C). Thus, these results support the idea that production of riboflavin intermediates by *S. epidermidis* is required for local MAIT cell activation and that *Staphylococcus* species generate riboflavin exclusively from GTP. To assess whether exposure to microbial antigens was sufficient to activate MAIT cells within the skin, we topically applied 5-OP-RU to SPF mice. Application of

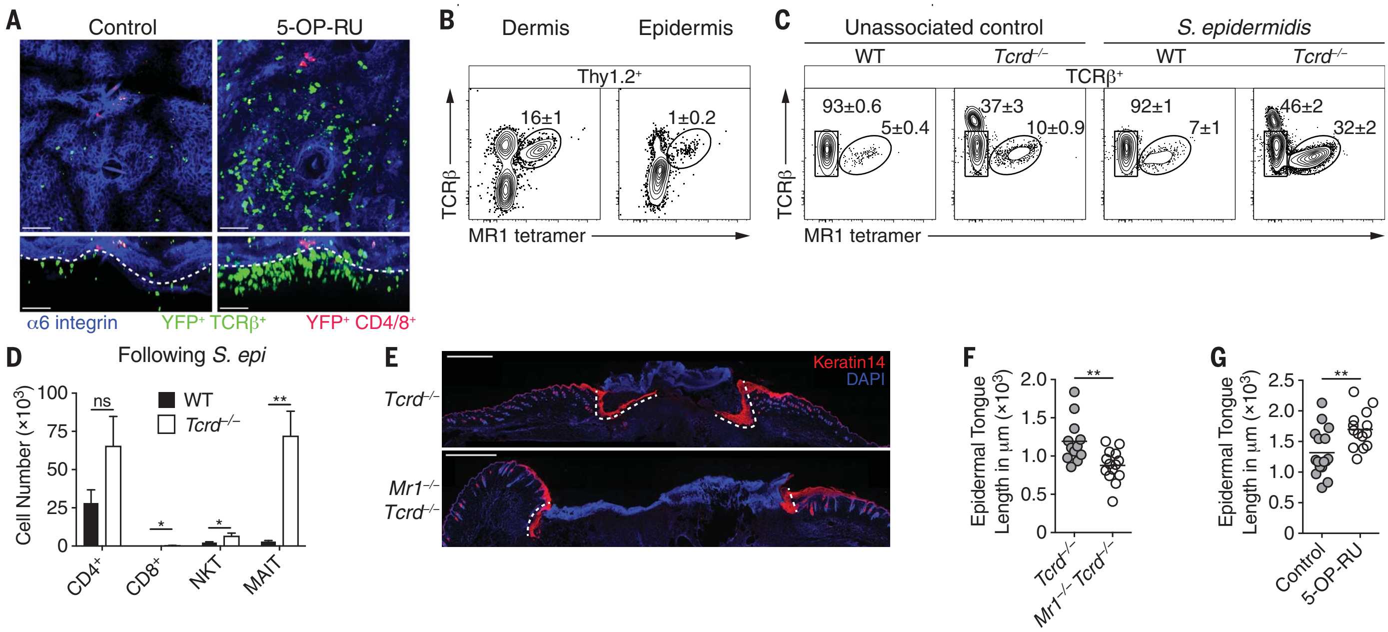


Fig. 5. MAIT cells promote tissue repair. (**A**) Representative confocal microscopy images of the skin of *Il17a*^{cre} *R26-STOP-YFP* mice that received topical 5-OP-RU and untreated controls. Shown is a projection along the *z* axis (top) or *x* axis (bottom). Colocalization of yellow fluorescent protein (YFP) and TCRβ is depicted in green, whereas the colocalization of YFP and either CD4 or CD8 is in red. White dashed lines denote the dermal-epidermal interface. Scale bars, 50 μm. (**B**) Flow cytometry of Thy1.2^{+} lymphocytes within the dermis and epidermis of WT mice. (**C**) Flow cytometry of TCRβ^{+} lymphocytes from the skin of *Tcrd*$^{-/-}$ and WT mice either associated with *S. epidermidis* LM061 or left untreated. The TCRβ^{hi} cells in the *Tcrd*$^{-/-}$ samples are DETCs that express αβ TCRs. (**D**) Number of cutaneous T cells in *Tcrd*$^{-/-}$ and WT mice after association with *S. epidermidis* LM061. (**E** and **F**) *S. epidermidis* LM061 was topically applied to the backs of MAIT cell–deficient *Mr1*$^{-/-}$ *Tcrd*$^{-/-}$ mice and *Tcrd*$^{-/-}$ littermates on days 0, 2, 4, and 6. Punch biopsies were taken through the back skin 12 days after the initial application, and animals were assessed 5 days later. (E) Representative immunofluorescence images of back skin wounds 5 days after punch biopsies. Tissue sections were immunolabeled with Keratin 14 (red), which stains the advancing epidermal tongues (demarcated with white dashed lines) during re-epithelialization of the wounds. Scale bars, 1 mm. (F) Quantification of the epidermal tongue length 5 days after wounding, with each dot representing the measured length of an individual epidermal tongue. (**G**) 5-OP-RU was topically applied to the backs of WT mice on days 0, 2, 4, and 6. Punch biopsies were taken through the back skin 12 days after the initial application, and animals were assessed 5 days later, when epidermal tongue lengths were quantified. Flow cytometry gate frequencies and graphs indicate means ± SEM. Data represent at least two experiments with four or more mice per group. **P* < 0.05 and ***P* < 0.01 as calculated with Student's *t* test. "ns" denotes that comparison was not significant.

antigen alone resulted in a significant increase of cutaneous IL-17A^{+} MAIT cells, with minimal effects on other T cells subsets (Fig. 4, D to H, and fig. S4, D and E). As with *S. epidermidis*, the MAIT cell response to 5-OP-RU resulted from local expansion (fig. S4, F and G), without an accompanying expansion of MAIT cells in other tissues (figs. S4, H to M).

To further interrogate whether antigen presentation is necessary for the MAIT cell response to skin commensals, we generated *Mr1*$^{f/f}$ conditional knockout mice by inserting LoxP sites flanking exon 3 (fig. S4N), resulting in the elimination of the α2 helix, which together with the α1 helix composes the ligand-binding cleft of MR1 (*15*). Although previous work has suggested that MAIT cells are selected on double-positive (DP) thymocytes similarly to NKT cells (*54*), this study used TCR transgenic T cells, which have been shown to display an unusual PLZF^{-} CD44^{-} phenotype (*26*). Therefore, whether thymocytes are necessary for the selection of WT MAIT cells remained to be determined. Analysis of *Mr1*$^{f/f}$ *Cd4*^{cre} mice revealed a profound reduction in MAIT cells compared with littermate controls, whereas other lineages were unaltered (fig. S4, O and P), indicating that MAIT cells are indeed selected on thymocytes.

Previous studies have suggested that B cells promote MAIT cell development and may present commensal antigens (*17*, *55*). However, *Mr1*$^{f/f}$ *Cd19*^{cre} mice and *Mr1*$^{f/f}$ littermates had similar numbers of MAIT cells before and after association with *S. epidermidis* (fig. S4Q). To determine whether the MAIT cell response to *S. epidermidis* required antigen recognition, we used a tamoxifen-inducible Cre recombinase to enable ablation of MR1 after MAIT cell development (*56*). Treating the resulting *Mr1*$^{f/f}$ *Rosa26*^{creERT2} mice and *Mr1*$^{f/f}$ littermates with tamoxifen did not affect MAIT cell numbers 3 weeks later, suggesting that continuous antigen presentation is not necessary to maintain MAIT cells in the skin under steady-state conditions (Fig. 4J). Tamoxifen treatment a week before the *S. epidermidis* association revealed that MR1-mediated antigen presentation is required for the MAIT cell response (Fig. 4, I and J, and fig. S4, R and S). Thus, the production of riboflavin metabolites by defined skin commensals and presentation of these antigens promotes local expansion and activation of cutaneous MAIT cells. Although cytokines have been shown to be sufficient to activate MAIT cells (*44*), our results support the idea that in the context of homeostatic responses to the microbiota, the MAIT cell response is strictly TCR-dependent.

MAIT cells promote tissue repair

To explore a potential role for MAIT cells in skin physiology, we first determined their localization within the skin. Because MAIT cells account for the vast majority of IL-17A–producing αβ T cells after topical application of 5-OP-RU (fig. S5A), we imaged the skin of *Il17a*^{cre} *R26-STOP-YFP* fate-mapping mice by means of confocal microscopy and found that MAIT cells predominantly resided near the interface of the dermis and epidermis, in close proximity to the basal layer (Fig. 5A). Dissociation of the dermis and epidermis confirmed that although

few MAIT cells are present in the epidermal layer, the majority are localized in the dermis (Fig. 5B). This localization contrasted with that of conventional $CD4^+$ or $CD8^+$ T cells, which localized close to hair follicles and within the epidermis, respectively (*38*). Together with the tissue remodeling signature of these cells (Figs. 2, C and J, and 3, Q and R), this supported the idea that MAIT cells may regulate tissue responses to injury.

Although alterations in MAIT cell frequencies have been associated with various pathologies, including skin conditions (*57*, *58*), the functions of these cells remain largely unclear. The previous difficulty assessing MAIT cell function in mice may have been due to the low frequency of MAIT cells in mouse tissues other than skin, as well as the potentially overlapping functions and compensatory responses of unconventional T cell subsets. NKT cell–deficient *Cd1d*$^{-/-}$ mice harbor greater numbers of thymic and splenic MAIT cells (*18*), suggesting that this subset of innate-like lymphocytes competes with the MAIT cell population. However, such a compensatory response was not observed in the skin of *Cd1d*$^{-/-}$ mice (fig. S5, B and C), likely because of the relatively small number of NKT cells in the skin compared with γδ T cells. Conversely, in the absence of γδ T cells, MAIT and NKT cell numbers were significantly increased at steady state and after *S. epidermidis* association (Fig. 5, C and D, and fig. S5, D to F). Therefore, γδ T cells restrict the MAIT cell population at steady state and constrain their response to commensals. This observation provided us with an approach to test the function of MAIT cells when competition with other innate-like lymphocytes is limited, a setting that is relevant to the lymphocyte composition of human skin, where γδ T cells are far less abundant than in mice (*59*).

Previous studies from our laboratory and others have indicated that innate-like lymphocytes can promote tissue repair. For instance, $CD8^+$ T cells that recognize commensal-derived N-formylated peptides presented by the MHC-Ib molecule H2-M3 accelerate wound closure in mice (*12*), whereas mouse dendritic epidermal T cells (DETCs) promote epidermal thickening and keratinocyte proliferation (*60*). We used an experimental model that involved taking a punch biopsy through the mouse back skin 12 days after the initial association with *S. epidermidis* and assessing the re-epithelialization of the wound 5 days later by measuring the epidermal tongue of proliferating keratinocytes (fig. S5G) (*12*, *61*). To minimize confounding effects of H2-M3–restricted $CD8^+$ T cells, we topically associated *Tcrd*$^{-/-}$ mice and MAIT cell–deficient *Mr1*$^{-/-}$ *Tcrd*$^{-/-}$ littermates with a strain of *S. epidermidis* that does not induce $CD8^+$ T cells (*12*). After wounding, the epidermal tongue lengths of MAIT cell–sufficient *Tcrd*$^{-/-}$ mice were significantly longer than those observed in MAIT cell–deficient *Mr1*$^{-/-}$ *Tcrd*$^{-/-}$ animals (Fig. 5, E and F), indicating that MAIT cells promote tissue repair in the absence of compensatory responses by γδ T cells.

To assess whether an increase in MAIT cells was sufficient to affect skin physiology and response to injury, we directly applied the 5-OP-RU antigen to the skin of WT mice before wounding to selectively increase MAIT cells (Fig. 4E). The topical application of 5-OP-RU significantly increased the epidermal tongue length (Fig. 5G), demonstrating that the microbially derived antigen 5-OP-RU was sufficient to promote wound healing in an intact host.

Discussion

After development, the skin harbors tissue-resident IL-$17A^+$ MAIT cells, and exposure to defined skin microbes or microbially derived antigens can further MAIT cell accumulation and strengthen the tissue repair functions of these cells in an antigen-dependent manner. The majority of skin commensals were capable of synthesizing riboflavin-derived metabolites (Fig. 4A), suggesting that these antigens contribute to skin health and homeostasis and may represent new therapeutic approaches for promoting tissue repair.

We also uncovered how early microbial encounters have long-term effects on the composition of the immune system. We show that MAIT cells are induced during a specific early-life window in response to riboflavin-synthesizing commensals, a phenomenon that permanently imprints the abundance of this subset in tissues. Previous work uncovered a role for early-life exposure to epithelial cell–derived butyrophilin-like proteins for $V\gamma7^+$ γδ intraepithelial lymphocyte development, gestational colonization for the induction of ILCs, and early colonization with the intestinal commensal *B. fragilis* to constrain the abundance of NKT cells in adults (*8*, *62*, *63*). Together, these findings imply that temporal windows of development and permanent imprinting of tissues may represent a general phenomenon for nonclassically restricted T cells. Our work also provides a possible explanation for the remarkable variation in MAIT cell abundance observed in humans (*21*, *22*). In contrast to mice, human T cells develop in utero, which supports the hypothesis that such early MAIT cell imprinting may occur before birth. Functionally mature MAIT cells have been found in second-trimester human fetuses (*64*). How, in humans, microbiota-derived riboflavin intermediates before and after birth contribute to the final maturation and long-term tissue imprinting of MAIT cells remains to be addressed. Our work proposes that a central strategy used by the immune system to interact with the microbiota may be through the recognition of canonical microbially derived antigens by unconventional T cells, including MAIT cells, a dialogue that may play a dominant role in the control of host physiology.

Materials and Methods

Mice

Il17a^{cre} (*65*), *Il23r*$^{gfp/gfp}$ (*66*), and *Mr1*$^{-/-}$ (*17*) mice have been previously described and were generously provided by B. Stockinger (Francis Crick Institute), M. Oukka, (Seattle Children's Research Institute), and S. Gilfillan (Washington University School of Medicine), respectively. Wild-type (WT) C57BL/6NTac and BALB/c SPF mice were purchased from Taconic Biosciences. CD45.1 (B6.SJL-Ptprc^{a} Pepc^{b}/BoyJ) and *Il1r1*$^{-/-}$ (C57BL/6-[KO]IL1r1) were obtained through the NIAID-Taconic exchange program. GF C57BL/6 mice were bred and maintained in the NIAID Microbiome Program gnotobiotic animal facility. *Cd1d*$^{-/-}$ (B6.129S6-Del(3Cd1d2-Cd1d1)1Sbp/J), *Cd4*^{cre} (Tg(Cd4-cre)1Cwi/BfluJ), *Cd19*^{cre} (B6.129P2(C)-Cd19$^{tm1(cre)Cgn}$/J), *Lta*$^{-/-}$ (B6.129S2-Lta^{tm1Dch}/J), *R26-STOP-YFP* (B6.129X1-Gt(ROSA)26Sor$^{tm1(EYFP)Cos}$/J), *R26*^{creERT2} (B6.129-Gt(ROSA)26Sor$^{tm1(cre/ERT2)Tyj}$/J), and *Tcrd*$^{-/-}$ (B6.129P2-Tcrd^{tm1Mom}/J) mice were purchased from The Jackson Laboratory. All mice were bred and maintained at an American Association for the Accreditation of Laboratory Animal Care (AAALAC)-accredited animal facility at the NIAID and housed in accordance with the procedures outlined in the Guide for the Care and Use of Laboratory Animals. All experiments were performed at NIAID under an animal study proposal (LISB-19E or LISB-20E) approved by the NIAID Animal Care and Use Committee. Unless otherwise noted, age- and sex-matched mice between 6 and 12 weeks of age were used for each experiment.

Generation of Mr1$^{f/f}$ mice

CRISPR guide RNAs (gRNAs) were synthesized following a published protocol (*67*). The following oligos were used to target the second and third introns of *Mr1*: 5′-TAGGTTAAAGC-CATCCTCCCAT-3′ and 5′-AAACATGGGAGGA-TGGCTTTAA-3′ (second) and 5′-TAGGCTAG-GCATGTTAAGAATG-3′ and 5′-AAACCATTCT-TAACATGCCTAG-3′ (third). Briefly, oligos were annealed and ligated into the pT7-gRNA vector (Addgene). These constructs were linearized and used as templates for in vitro transcription (IVT) using the MEGAshortscript T7 Transcription Kit (Thermo Fisher Scientific), which yielded the gRNAs. Cas9 mRNA was generated by using linearized pT3TS-nCas9n (Addgene) as the template for IVT with the mMESSAGE mMACHINE T3 Transcription Kit (Thermo Fisher Scientific). Both the Cas9 mRNA and the gRNAs were purified using the MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific) and eluted in RNase-free water. C57BL/6NCr embryos were microinjected with Cas9 mRNA (100 ng/μL), both gRNAs (100 ng/μL), and the following two PAGE-purified ssDNA

oligos (300 ng/μL; Integrated DNA Technologies), which contain the LoxP sequences flanked by 75 bases of homology: 5′-CAAAAGTGCTTTCAGCCATGCTCCAGTTGTACCACTTGAGAAGTCTCTCTCTGTCTCTGTTTAAAGCCATCCTCCATAACTTCGTATAATGTATGCTATACGAAGTTATCATAGGCTGTATATGTTTCAGGGCTTCACACCTACCAGAGAATGATTGGCTGTGAGTTGCTAGAAGATGGCAGCA-3′ and 5′-TGGCTAAAGAGGTTCTTGGAATATGGAAGAGATACCCTAGAAAGAACAGGTAATGGGAAGAGAGAACACCTCATATAACTTCGTATAATGTATGCTATACGAAGTTATTCTTAACATGCCTAGGAAAACTGGCTTTGTGTCTGACTTCCTCTTCTGCAGGTGATATTCGCCCTGTGCTGAAAT-3′ and surgically transferred to recipient females. Progeny were screened by polymerase chain reaction (PCR) amplification with primers 5′-TTTCTGGCTCCTGCGTATCT-3′ and 5′-ATTTCAGCACAGGGCGAATA-3′, followed by Sanger sequencing with the primer 5′-CTAATTCCCCAAGGCAAGC-3′. The founder was bred to C57BL/6NTac mice to propagate the strain.

Inducible deletion of MR1

Mr1$^{f/f}$ R26^{creERT2} and *Mr1$^{f/f}$* littermate controls were injected intraperitoneally with 3 mg of tamoxifen (Sigma-Aldrich) in a corn oil-ethanol (95:5) mixture 8, 6, 4, and 2 days before association with *S. epidermidis* NIHLM061.

Commensal culture and colonization

For the isolation of intestinal commensals, a stool pellet was homogenized in 200 μL of BHI and 10 μL of the homogenate was streaked on MacConkey (BD Biosciences) and de Man, Rogosa, and Sharpe (MRS; Sigma-Aldrich) agar plates and incubated at 37°C under aerobic and microaerophilic conditions, respectively. After 48 hours of incubation, colonies bearing distinct morphologies were sub-cultured to obtain pure cultures. Additionally, reiterative 1:10 dilutions in Brain Heart Infusion (BHI) broth allowed isolation of single bacteria in 96-well plates under aerobic conditions. Taxonomic identification of pure isolates was determined with Sanger sequencing. To prepare bacterial cultures for oral gavage, 10 μL of each frozen strain was propagated on Columbia agar with 5% sheep blood (Thermo Scientific) and incubated under microaerophilic conditions overnight at 36°C. Two to three colonies of each strain were inoculated in Peptone Yeast Glucose (PYG) broth and incubated at 37°C under microaerophilic conditions for 12 hours, except for *Lactobacillus* strains, which were incubated for 24 hours. All cultures were normalized to an optical density (OD) of 1 at 600 nm, mixed in equal volumes when required (5-species community), centrifuged, and resuspended in 10% the volume of PYG broth. Neonates received 20 μL of the bacterial culture by oral gavage, whereas adult animals received 200 μL. For the topical association with *S. epidermidis* NIH05001, NIHLM061, and NIHLM087 (referred to as 05001, LM061, and LM087 in the text) (68), strains were cultured for 18 hours in Tryptic Soy Broth (TSB) at 37°C, to a density of approximately 10^9 colony-forming units (CFU)/mL. Mice were colonized by topically applying up to 5 mL of culture across the entire skin surface using a sterile swab. Association was repeated every other day, for a total of four times. To quantitate *S. epidermidis* colonization, the ear skin was swabbed with a sterile cotton swab soaked in TSB. Swabs were streaked on blood agar and plates were then incubated at 37°C under aerobic conditions for 18 hours before counting CFUs.

Generation of S. epidermidis mutant

The *S. epidermidis* NIHLM087 Δ*ribD* mutant was constructed by using a recently described method (Chen *et al.*, submitted). Approximately 1 kb up- and downstream of the gene *ribD* were amplified with the following primers: 5′-CGGTATCGATAAGCTTGATATCTGGAATAGTTTGAGCGTTAA-3′ and 5′-GATACCTGTAAACATAGACATTCTACTCAATTGATCACCTC-3′ and 5′-GAGGTGATCAATTGAGTAGAATGTCTATGTTTACAGGTATC-3′ and 5′-TATAGGGCGAATTGGAGCTCTTACGTGTTAACACACCATT-3′. These two flanking regions were assembled into the plasmid pIMAY (Addgene) (69) by using Gibson Assembly (New England Biolabs) (70). The resulting plasmid was transformed into *S. epidermidis* NIHLM087. Temperature shift to 37°C and anhydrotetracycline were then used to select for allelic recombination events that resulted in a *ribD* deletion without a genomic scar. Successful knockouts were confirmed by means of colony PCR and sequencing.

Parabiosis experiments and surgery

Age-matched CD45.1 and CD45.2 mice were topically associated with *S. epidermidis* NIHLM087 and 7 weeks after the initial colonization, parabiosis surgery was performed as described (71). Animals were treated with oral antibiotics for 2 weeks and remained conjoined for 13 weeks before analysis.

In vivo cytokine blockade

Mice were injected subcutaneously with 1 mg of either anti-IL-23R antibody (21A4; Merck) or mIgG1 isotype control (27F11; Merck) 2 d before the initial application of *S. epidermidis* NIHLM061 and again on day 6. Anti-IL-18 monoclonal antibody (SK113AE-4) (72) was purified from hybridoma supernatant using Protein G Sepharose beads (Abcam). Mice were injected intraperitoneally with either 1 mg of anti-IL-18 or saline (vehicle) 2 days before the initial application of *S. epidermidis* NIHLM061 and again on days 1, 5, 8, and 11.

Human skin specimens

Normal adult human skin was obtained from five patients at the University of California, San Francisco, undergoing elective surgery, in which healthy skin was discarded as a routine procedure. The study was conducted in accordance with the Declaration of Helsinki principles.

Human skin processing

Skin samples were stored at 4°C in a sterile container with phosphate-buffered saline (PBS) and gauze until the time of digestion. Subcutaneous fat was removed, skin was minced finely with dissection scissors, and mixed in a six-well plate with 3 mL of digestion buffer consisting of 0.8 mg/mL collagenase type 4 (Worthington), 0.02 mg/mL DNase I (Sigma-Aldrich), 10% fetal bovine serum (FBS), 1% hydroxyethylpiperazine ethane sulfonic acid (HEPES), and 1% penicillin/streptavidin in RPMI medium. Samples were incubated overnight in 5% CO_2 and filtered through a 100-μm filter with wash buffer (2% FBS, 1% penicillin/streptavidin in RPMI medium).

Mouse tissue processing

To isolate skin cells, ear pinnae were excised and separated into ventral and dorsal sheets, which were digested by placing them dermal side down in RPMI 1640 media supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 1 mM nonessential amino acids, 50 mM β-mercaptoethanol, 20 mM HEPES, 100 U/mL penicillin, 100 mg/mL streptomycin, 0.5 mg/mL DNase I (Sigma-Aldrich), and 0.25 mg/ml of Liberase TL purified enzyme blend (Roche), and incubated for an hour and 45 min at 37°C. Digested skin sheets were homogenized by using the Medimachine tissue homogenizer system (BD Biosciences) and passed through 50-μm filters. To separate the epidermis from the dermis, ear pinnae were first digested for 45 min with 500 CU Dispase (Corning Life Sciences), and curved forceps were used to remove the epidermis, before proceeding with the Liberase TL digestion as previously described. Lungs were diced and incubated in 2 mL of pre-warmed RPMI 1640 containing 1 mg/mL DNase I (Sigma-Aldrich) and 0.25 mg/ml of Liberase TL (Roche) for 45 min in a 37°C water bath. The digested tissue was passed through a 70-μm filter, centrifuged, resuspended in 5 mL of 40% Percoll (Sigma-Aldrich), and centrifuged again. Thymus, spleen, and liver were dissociated through 70-μm filters, after which 40% Percoll (Sigma-Aldrich) was used to remove hepatocytes from liver samples. Small intestine lamina propria were processed as previously described (73).

In vitro restimulation

Cells were cultured for 2.5 hours at 37°C in RPMI complete media (RPMI 1640 supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 1 mM nonessential amino acids, 50 mM β-mercaptoethanol, 20 mM HEPES, 100 U/mL penicillin, and 100 mg/mL

streptomycin) containing 50 ng/mL of phorbol myristate acetate (PMA; Sigma-Aldrich), 5 mg/mL ionomycin (Sigma-Aldrich), and a 1:1000 dilution of GolgiPlug (BD Biosciences).

Flow cytometric analysis

Fluorophore-conjugated antibodies that were used are listed in table S1. 5-OP-RU and 6-FP–loaded mMR1 and hMR1 tetramers and PBS57-loaded mCD1d tetramers were provided by the NIH Tetramer Core Facility and cells were stained in RPMI complete media for 1 hour at room temperature. For intracellular cytokine staining, cells were fixed and permeabilized using the Fixation/Permeabilization Solution Kit (BD Biosciences) and stained with fluorophore-conjugated antibodies for at least 60 min at 4°C. For transcription factor staining, cells were fixed and permeabilized with the Foxp3/Transcription Factor Staining Buffer Set (eBioscience) and stained with fluorophore-conjugated antibodies for at least 60 min at 4°C. All staining was performed in the presence of purified anti-mouse CD16/32 and purified rat gamma globulin (Jackson ImmunoResearch). Dead cells were excluded from live samples using 4′,6-diamidino-2-phenylindol (DAPI; Sigma-Aldrich), whereas either LIVE/DEAD Fixable Blue Dead Cell Stain Kit (Invitrogen Life Technologies) or Ghost Dye Violet 510 (Tonbo Biosciences) was used in fixed samples. All antibodies were purchased from BD Biosciences, BioLegend, or eBioscience. Cells were acquired on an LSRII flow cytometer (BD Biosciences) running FACSDiva software (BD Biosciences), and data were analyzed by using FlowJo (TreeStar). Unless otherwise indicated, mouse MAIT cells were gated as MR1 tetramer$^+$ TCRβ$^+$ TCRγδ$^-$ CD3$^+$ CD90.2$^+$ CD45$^+$ (CD44$^+$ as well in tissues other than skin); NKT cells as CD1d tetramer$^+$ TCRβ$^+$ TCRγδ$^-$ CD3$^+$ CD90.2$^+$ CD45$^+$; CD4$^+$ T cells as CD4$^+$ CD8β$^-$ MR1 tetramer$^-$ CD1d tetramer$^-$ TCRβ$^+$ TCRγδ$^-$ CD3$^+$ CD90.2$^+$ CD45$^+$; CD8$^+$ T cells as CD8β$^+$ CD4$^-$ MR1 tetramer$^-$ CD1d tetramer$^-$ TCRβ$^+$ TCRγδ$^-$ CD3$^+$ CD90.2$^+$ CD45$^+$; and γδ T cells as TCRγδ$^+$ TCRβ$^-$ CD3$^+$ CD90.2$^+$ CD45$^+$.

Tetramer enrichment

For the enrichment of MAIT cells, samples were stained with phycoerythrin (PE)-conjugated MR1 tetramer, bound to anti-PE MicroBeads (Miltenyi Biotec), and enriched on an autoMACS Pro Separator (Miltenyi Biotec) using the POSSEL_S program.

Confocal microscopy

Ear pinnae were split with forceps, fixed in 1% paraformaldehyde (Electron Microscopy Sciences) overnight at 4°C, and blocked in PBS containing 1% bovine serum albumin (BSA) and 0.25% Triton X-100 for 2 hours at room temperature. Tissues were first stained with anti-CD4 (RM4-5, eBioscience), anti-CD8α (53-6.7, eBioscience), anti-CD49f (GoH3, eBioscience), anti-GFP (A21311, Life Technologies), and anti-TCRβ (H57-597, eBioscience) antibodies overnight at 4°C, washed three times with PBS, and then stained with DAPI (Sigma-Aldrich) overnight before being mounted with ProLong Gold (Invitrogen). Ear pinnae images were captured on a Leica TCS SP8 confocal microscope equipped with HyD and PMT detectors and a 40X oil objective (HC PL APO 40X/1.3 oil). Images were analyzed using Imaris (Bitplane).

Wounding and epifluorescence microscopy of wound tissue

Wounding and quantitation of wound healing were performed as previously described (*61*). Male mice in the telogen phase of the hair cycle were anesthetized with ketamine/ xylazine and their back skin was then shaved using a ChroMini Pro (Moser). A 6-mm biopsy punch was first used to partially perforate the skin and then iris scissors were used to cut the epidermis and dermis along the punch outline, resulting in a circular, full-thickness wound. 5 d after wounding, the skin tissue was excised and fixed in 4% paraformaldehyde for 4 hours at 4°C, incubated overnight in 30% sucrose at 4°C, embedded in OCT compound (Tissue-Tek), frozen on dry ice, and cryo-sectioned (20-μm-thick sections). Sections were fixed in 4% paraformaldehyde for 10 min at room temperature, rinsed with PBS, permeabilized with 0.1% Triton X-100 (Sigma) for 10 min, and blocked for 1 hour at room temperature in blocking buffer (2.5% normal goat serum, 1% BSA, 0.3% Triton X-100 in PBS). Chicken anti-mouse Keratin 14 antibody (Poly9060, BioLegend) was diluted at 1:400 in blocking buffer containing rat gamma globulin and anti-CD16/32 and incubated overnight at 4°C. After washing with PBS, sections were stained with a polyclonal anti-chicken IgY-AlexaFluor647 secondary antibody (Jackson ImmunoResearch) at a 1:800 dilution for 1 hour at room temperature, washed with PBS, stained with DAPI, washed with PBS, and mounted with ProLong Gold (Invitrogen). Wound images were captured with a Leica DMI 6000 widefield epifluorescence microscope equipped with a Leica DFC360X monochrome camera. Tiled and stitched images of wounds were collected using a 20X/0.4 NA dry objective. Images were analyzed using Imaris (Bitplane).

16S rRNA gene sequencing and analysis

Fecal DNA was purified using the MagAttract PowerMicrobiome DNA/RNA kit (Qiagen). Amplification of the V4 hypervariable region of the bacterial 16*S* rRNA gene was performed using the 515f and 806r primers (515F: 5′-GTGCCAGCMGCCGCGGTAA-3′; 806R: 5′-GGACTACHVGGGTWTCTAAT-3′), followed by an additional PCR to append different barcodes to each sample. Amplicons were quantified using Kapa Library Quantification Complete Kit (ROX Low) (Kapa Biosystems) and pooled at equimolar concentrations before being sequenced on a MiSeq (Illumina) using the v3 MiSeq Reagent kit (Illumina). The DADA2 algorithm (*74*) was used to denoise raw 16*S* reads after primer trimming and to tabulate sequence variants. Read counts were converted to relative abundance and taxonomy was assigned to each sequence variant using the RDP Training Set 16. Sequence variants were collapsed by family-level classifications, and families present in greater than 5% abundance in at least one time point were graphically represented, whereas remaining families were collapsed into the category "Other."

Bacterial whole-genome sequencing

For the *E. faecalis*, *K. oxytoca*, *L. johnsonii*, *L. murinus*, and *P. mirabilis* isolates, DNA libraries were prepared by using the paired-ended Nextera DNA Flex Library Prep Kit (Illumina). Sequencing was performed on a NextSeq 500 (Illumina) using the High Output v2 Kit (Illumina).

Assessing the presence of riboflavin synthesis genes

Protein sequences of representative riboflavin biosynthesis genes were downloaded from GenBank: *ribA* (NP_764994), *ribB* (CDO12415), *ribD* (NP_764996), *ribE* (NP_764995), *ribH* (NP_764993), and *pyrp2* (NP_764225). Isolates were either sequenced as described above or genomes were downloaded from National Center for Biotechnology Information (NCBI): *Candidatus Arthromitus* (BioProject PRJDA67835), *S. epidermidis* NIH05001 (BioProject PRJNA62391), *S. epidermidis* NIHLM061 (BioProject PRJNA62355), and *S. epidermidis* NIHLM087 (BioProject PRJNA62349). For taxa that are generally present in the intestinal and skin microbiota, genomes of different strains were downloaded from NCBI: Corynebacteriaceae (txid 1653), Propionibacteriaceae (txid 31957), Staphylococcaceae (txid 90964), Streptococcaceae (txid 1300), Bacteroidaceae (txid 815), Bifidobacteriaceae (txid 31953), Clostridiaceae (txid 31979), Enterobacteriaceae (txid 543), Enterococcaceae (txid 81852), Eubacteriaceae (txid 186806), Hungateiclostridiaceae (txid 2304686), Lachnospiraceae (txid 186803), Lactobacillaceae (txid 33958), Peptococcaceae (txid 186807), Peptostreptococcaceae (txid 186804), Ruminococcaceae (txid 541000), and Tannerellaceae (txid 2005525). These nucleotide scaffolds were used as subjects in a tBLASTn search using representative riboflavin biosynthesis genes, using an e-value $<1\times10^{-10}$. Alignments that met this threshold were considered as homologs, and thus those genomes were considered to contain that gene.

RNA-seq and analysis

The following populations were isolated by means of fluorescence-activated cell sorting

(FACS) from the ear pinnae of C57BL/6NTac mice associated with *S. epidermidis* NIHLM061 and untreated control animals: MAIT cells (MR1 tetramer$^+$ TCRβ$^+$ TCRγδ $^-$ CD90.2$^+$ CD45$^+$ DAPI$^-$) and CD4$^+$ T cells (CD4$^+$ CD25$^-$ CD8β$^-$ MR1 tetramer$^-$ TCRβ$^+$ TCRγδ $^-$ CD90.2$^+$ CD45$^+$ DAPI$^-$). RNA was isolated using the Arcturus PicoPure RNA Isolation Kit (Applied Biosystems), cDNA was synthesized using the SMART-Seq v4 Ultra Low Input RNA Kit (Clontech Laboratories), libraries were generated using Nextera XT DNA Library Preparation Kit (Illumina), and the samples were sequenced on a NextSeq 500 (Illumina) using the High Output v2 Kit (Illumina). Fastq files were mapped with STAR (*75*) to the mouse mm10 reference genome with default parameters. For analysis, the Hypergeometric Optimization of Motif EnRichment (HOMER) program was used (*76*). Tag directories were created with default parameters. To analyze differential gene expression, reads per gene were counted for two replicates per condition with analyzeRepeats rna with parameters -noadj and -condenseGenes. Subsequentially, getDiffExpression.pl was used with default parameters [log2 fold change 1, false discovery rate (FDR) < 5%] using DESeq2 and rlog normalization. For GO term enrichment analysis Metascape (*77*) was used.

Single-cell RNA-seq and analysis

Lymphocytes from the ear pinnae (skin), spleens, lungs, and livers of pooled WT mice were stained with antibodies against surface markers and TotalSeq-A hashtag oligonucleotide (HTO) antibodies (BioLegend; hashtag 1 for skin, 2 for livers, 3 for lungs, and 4 for spleens). MAIT cells were sorted from each organ as MR1 tetramer$^+$ TCRβ$^+$ TCRγδ$^-$ CD3$^+$ CD90.2$^+$ B220$^-$ CD11b$^-$ CD11c$^-$ CD49f$^-$ CD45$^+$ (10,000 cells from skin, spleens, and lungs and 3500 cells from livers). Samples were pooled, encapsulated into droplets using the Chromium Single Cell Controller (10X Genomics), and libraries were prepared using Chromium Single Cell 3′ Reagent Kits v2 (10X Genomics). The mRNA library was prepared following the 10X Genomics user guide, while the HTO library was prepared according to published guidelines (*78*). Libraries were sequenced on a NextSeq 500 with 10% of the lane occupied by the HTO library and the other 90% by the mRNA library.

Data was demultiplexed by using Cell Ranger 3.0.1 resulting in data for 6463 cells. HTO libraries were processed by using CITE-seq-Count (DOI: 10.5281/zenodo.2590196). Subsequently, data was analyzed using Seurat 3.0 (*79*). Cells with less than 900 or more than 3000 gene counts, more than 5% mitochondrial contamination, or more than 10,000 RNA counts were filtered out, leaving 5678 cells for downstream analysis. Data was normalized, scaled, principal components analysis was performed, and neighbors were found using five dimensions. Uniform manifold approximation and projection (UMAP) reduction was performed on this dataset with 10 dimensions. HTO were read in and assigned to the single cells and only singlets were used for further analysis leaving 585 cells. To define origins of cells, number of HTOs per cluster were considered and classified as "Skin," "Spleen/Lung/Liver," or not assigned ("NA"). To analyze gene expression between tissues, all cells, including cells without HTO from the respective clusters were used.

Statistical analysis

Two-tailed unpaired Student's *t* test was performed by using Prism 7 (GraphPad) to determine statistical significance. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$. "ns" denotes that the comparison was not significant.

REFERENCES AND NOTES

1. Y. Belkaid, T. W. Hand, Role of the microbiota in immunity and inflammation. *Cell* **157**, 121–141 (2014). doi: 10.1016/j.cell.2014.03.011; pmid: 24679531
2. M. G. Rooks, W. S. Garrett, Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341–352 (2016). doi: 10.1038/nri.2016.42; pmid: 27231050
3. E. S. Gollwitzer *et al*., Lung microbiota promotes tolerance to allergens in neonates via PD-L1. *Nat. Med.* **20**, 642–647 (2014). doi: 10.1038/nm.3568; pmid: 24813249
4. T. C. Scharschmidt *et al*., A wave of regulatory T cells into neonatal skin mediates tolerance to commensal microbes. *Immunity* **43**, 1011–1021 (2015). doi: 10.1016/j.immuni.2015.10.016; pmid: 26588783
5. K. Mao *et al*., Innate and adaptive lymphocytes sequentially shape the gut microbiota and lipid metabolism. *Nature* **554**, 255–259 (2018). doi: 10.1038/nature25437; pmid: 29364878
6. Z. Al Nabhani *et al*., A weaning reaction to microbiota is required for resistance to immunopathologies in the adult. *Immunity* **50**, 1276–1288.e5 (2019). doi: 10.1016/j.immuni.2019.02.014; pmid: 30902637
7. J. Cahenzli, Y. Köller, M. Wyss, M. B. Geuking, K. D. McCoy, Intestinal microbial diversity during early-life colonization shapes long-term IgE levels. *Cell Host Microbe* **14**, 559–570 (2013). doi: 10.1016/j.chom.2013.10.004; pmid: 24237701
8. T. Olszak *et al*., Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**, 489–493 (2012). doi: 10.1126/science.1219328; pmid: 22442383
9. D. An *et al*., Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* **156**, 123–133 (2014). doi: 10.1016/j.cell.2013.11.042; pmid: 24439373
10. M. G. Constantinides, Interactions between the microbiota and innate and innate-like lymphocytes. *J. Leukoc. Biol.* **103**, 409–419 (2018). doi: 10.1002/JLB.3RI0917-378R; pmid: 29345366
11. V. K. Ridaura *et al*., Contextual control of skin immunity and inflammation by *Corynebacterium*. *J. Exp. Med.* **215**, 785–799 (2018). doi: 10.1084/jem.20171079; pmid: 29382696
12. J. L. Linehan *et al*., Non-classical immunity controls microbiota impact on skin immunity and tissue repair. *Cell* **172**, 784–796.e18 (2018). doi: 10.1016/j.cell.2017.12.033; pmid: 29358051
13. A. S. Ismail *et al*., Gammadelta intraepithelial lymphocytes are essential mediators of host-microbial homeostasis at the intestinal mucosal surface. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8743–8748 (2011). doi: 10.1073/pnas.1019574108; pmid: 21555560
14. C. Jin *et al*., Commensal microbiota promote lung cancer development via γδ T cells. *Cell* **176**, 998–1013.e16 (2019). doi: 10.1016/j.cell.2018.12.040; pmid: 30712876
15. L. Kjer-Nielsen *et al*., MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* **491**, 717–723 (2012). doi: 10.1038/nature11605; pmid: 23051753
16. L. Le Bourhis *et al*., Antimicrobial activity of mucosal-associated invariant T cells. *Nat. Immunol.* **11**, 701–708 (2010). doi: 10.1038/ni.1890; pmid: 20581831
17. E. Treiner *et al*., Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature* **422**, 164–169 (2003). doi: 10.1038/nature01433; pmid: 12634786
18. H. F. Koay *et al*., A three-stage intrathymic development pathway for the mucosal-associated invariant T cell lineage. *Nat. Immunol.* **17**, 1300–1311 (2016). doi: 10.1038/ni.3565; pmid: 27668799
19. M. Dusseaux *et al*., Human MAIT cells are xenobiotic-resistant, tissue-targeted, CD161hi IL-17-secreting T cells. *Blood* **117**, 1250–1259 (2011). doi: 10.1182/blood-2010-08-303339; pmid: 21084709
20. C. K. Vorkas *et al*., Mucosal-associated invariant and γδ T cell subsets respond to initial *Mycobacterium tuberculosis* infection. *JCI Insight* **3**, e121899 (2018). doi: 10.1172/jci.insight.121899; pmid: 30282828
21. G. Ben Youssef *et al*., Ontogeny of human mucosal-associated invariant T cells and related T cell subsets. *J. Exp. Med.* **215**, 459–479 (2018). doi: 10.1084/jem.20171739; pmid: 29339446
22. P. Chen *et al*., Circulating mucosal-associated invariant T cells in a large cohort of healthy Chinese individuals from newborn to elderly. *Front. Immunol.* **10**, 260 (2019). doi: 10.3389/fimmu.2019.00260; pmid: 30838000
23. J. E. Ussher, P. Klenerman, C. B. Willberg, Mucosal-associated invariant T-cells: New players in anti-bacterial immunity. *Front. Immunol.* **5**, 450 (2014). doi: 10.3389/fimmu.2014.00450; pmid: 25339949
24. H. Wang *et al*., MAIT cells protect against pulmonary Legionella longbeachae infection. *Nat. Commun.* **9**, 3350 (2018). doi: 10.1038/s41467-018-05202-8; pmid: 30135490
25. B. van Wilgenburg *et al*., MAIT cells contribute to protection against lethal influenza infection in vivo. *Nat. Commun.* **9**, 4706 (2018). doi: 10.1038/s41467-018-07207-9; pmid: 30413689
26. A. Rahimpour *et al*., Identification of phenotypically and functionally heterogeneous mouse mucosal-associated invariant T cells using MR1 tetramers. *J. Exp. Med.* **212**, 1095–1108 (2015). doi: 10.1084/jem.20142110; pmid: 26101265
27. A. J. Corbett *et al*., T-cell activation by transitory neo-antigens derived from distinct microbial pathways. *Nature* **509**, 361–365 (2014). doi: 10.1038/nature13160; pmid: 24695216
28. Y. G. Kim *et al*., Neonatal acquisition of *Clostridia* species protects against colonization by bacterial pathogens. *Science* **356**, 315–319 (2017). doi: 10.1126/science.aag2029; pmid: 28428425
29. I. I. Ivanov *et al*., Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009). doi: 10.1016/j.cell.2009.09.033; pmid: 19836068
30. V. Gaboriau-Routhiau *et al*., The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity* **31**, 677–689 (2009). doi: 10.1016/j.immuni.2009.08.020; pmid: 19833089
31. F. Legoux *et al*., Microbial metabolites control the thymic development of mucosal-associated invariant T cells. *Science* eaaw2719 (2019). doi: 10.1126/science.aaw2719; pmid: 31467190
32. D. M. Chu *et al*., Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat. Med.* **23**, 314–326 (2017). doi: 10.1038/nm.4272; pmid: 28112736
33. J. L. Combellick *et al*., Differences in the fecal microbiota of neonates born at home or in the hospital. *Sci. Rep.* **8**, 15660 (2018). doi: 10.1038/s41598-018-33995-7; pmid: 30353125
34. K. Li *et al*., Synthesis, stabilization, and characterization of the MR1 ligand precursor 5-amino-6-D-ribitylaminouracil (5-A-RU). *PLOS ONE* **13**, e0191837 (2018). doi: 10.1371/journal.pone.0191837; pmid: 29401462
35. T. Sano *et al*., An IL-23R/IL-22 circuit regulates epithelial serum amyloid A to promote local effector Th17 responses. *Cell* **163**, 381–393 (2015). doi: 10.1016/j.cell.2015.08.061; pmid: 26411290
36. J. Dias *et al*., The CD4$^-$CD8$^-$ MAIT cell subpopulation is a functionally distinct subset developmentally related to the main CD8$^+$ MAIT cell pool. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11513–E11522 (2018). doi: 10.1073/pnas.1812273115; pmid: 30442667
37. A. L. Byrd, Y. Belkaid, J. A. Segre, The human skin microbiome. *Nat. Rev. Microbiol.* **16**, 143–155 (2018). doi: 10.1038/nrmicro.2017.157; pmid: 29332945
38. S. Naik *et al*., Commensal-dendritic-cell interaction specifies a unique protective skin immune signature. *Nature* **520**, 104–108 (2015). doi: 10.1038/nature14052; pmid: 25539086

39. P. De Togni et al., Abnormal development of peripheral lymphoid organs in mice deficient in lymphotoxin. *Science* **264**, 703–707 (1994). doi: 10.1126/science.8171322; pmid: 8171322
40. X. Jiang et al., Dermal γδ T cells do not freely re-circulate out of skin and produce IL-17 to promote neutrophil infiltration during primary contact hypersensitivity. *PLOS ONE* **12**, e0169397 (2017). doi: 10.1371/journal.pone.0169397; pmid: 28081153
41. S. Y. Thomas et al., PLZF induces an intravascular surveillance program mediated by long-lived LFA-1-ICAM-1 interactions. *J. Exp. Med.* **208**, 1179–1188 (2011). doi: 10.1084/jem.20102630; pmid: 21624939
42. M. Salou et al., A common transcriptomic program acquired in the thymus defines tissue residency of MAIT and NKT subsets. *J. Exp. Med.* **216**, 133–151 (2019). doi: 10.1084/jem.20181483; pmid: 30518599
43. M. J. Sobkowiak et al., Tissue-resident MAIT cell populations in human oral mucosa exhibit an activated profile and produce IL-17. *Eur. J. Immunol.* **49**, 133–143 (2019). doi: 10.1002/eji.201847759; pmid: 30372518
44. B. van Wilgenburg et al., MAIT cells are activated during human viral infections. *Nat. Commun.* **7**, 11653 (2016). doi: 10.1038/ncomms11653; pmid: 27337592
45. S. Naik et al., Compartmentalized control of skin immunity by resident commensals. *Science* **337**, 1115–1119 (2012). doi: 10.1126/science.1225152; pmid: 22837383
46. M. P. Seiler et al., Elevated and sustained expression of the transcription factors Egr1 and Egr2 controls NKT lineage differentiation in response to TCR signaling. *Nat. Immunol.* **13**, 264–271 (2012). doi: 10.1038/ni.2230; pmid: 22306690
47. F. Bussolino et al., Hepatocyte growth factor is a potent angiogenic factor which stimulates endothelial cell motility and growth. *J. Cell Biol.* **119**, 629–641 (1992). doi: 10.1083/jcb.119.3.629; pmid: 1383237
48. M. Miyake, S. Goodison, A. Lawton, E. Gomes-Giacoia, C. J. Rosser, Angiogenin promotes tumoral growth and angiogenesis by regulating matrix metallopeptidase-2 expression via the ERK1/2 pathway. *Oncogene* **34**, 890–901 (2015). doi: 10.1038/onc.2014.2; pmid: 24561529
49. A. Toulon et al., A role for human skin-resident T cells in wound healing. *J. Exp. Med.* **206**, 743–750 (2009). doi: 10.1084/jem.20081787; pmid: 19307328
50. Z. He, C. H. Ong, J. Halper, A. Bateman, Progranulin is a mediator of the wound response. *Nat. Med.* **9**, 225–229 (2003). doi: 10.1038/nm816; pmid: 12524533
51. S. LeBlanc et al., CEACAM1 deficiency delays important wound healing processes. *Wound Repair Regen.* **19**, 745–752 (2011). doi: 10.1111/j.1524-475X.2011.00742.x; pmid: 22092845
52. M. Demaria et al., An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev. Cell* **31**, 722–733 (2014). doi: 10.1016/j.devcel.2014.11.012; pmid: 25499914
53. A. Grochot-Przeczek et al., Heme oxygenase-1 accelerates cutaneous wound healing in mice. *PLOS ONE* **4**, e5803 (2009). doi: 10.1371/journal.pone.0005803; pmid: 19495412
54. N. Seach et al., Double-positive thymocytes select mucosal-associated invariant T cells. *J. Immunol.* **191**, 6002–6009 (2013). doi: 10.4049/jimmunol.1301212; pmid: 24244014
55. E. Martin et al., Stepwise development of MAIT cells in mouse and human. *PLOS Biol.* **7**, e54 (2009). doi: 10.1371/journal.pbio.1000054; pmid: 19278296
56. A. Ventura et al., Restoration of p53 function leads to tumour regression in vivo. *Nature* **445**, 661–665 (2007). doi: 10.1038/nature05541; pmid: 17251932
57. M. B. M. Teunissen et al., The IL-17A-producing CD8$^+$ T-cell population in psoriatic lesional skin comprises mucosa-associated invariant T cells and conventional T cells. *J. Invest. Dermatol.* **134**, 2898–2907 (2014). doi: 10.1038/jid.2014.261; pmid: 24945094
58. J. Li et al., The frequency of mucosal-associated invariant T cells is selectively increased in dermatitis herpetiformis. *Australas. J. Dermatol.* **58**, 200–204 (2017). doi: 10.1111/ajd.12456; pmid: 26940855
59. J. Suwanpradid, Z. E. Holcomb, A. S. MacLeod, Emerging skin T-cell functions in response to environmental insults. *J. Invest. Dermatol.* **137**, 288–294 (2017). doi: 10.1016/j.jid.2016.08.013; pmid: 27784595
60. J. Jameson et al., A role for skin γδ T cells in wound repair. *Science* **296**, 747–749 (2002). doi: 10.1126/science.1069639; pmid: 11976459
61. B. E. Keyes et al., Impaired epidermal to dendritic T cell signaling slows wound repair in aged skin. *Cell* **167**, 1323–1338.e14 (2016). doi: 10.1016/j.cell.2016.10.052; pmid: 27863246
62. R. Di Marco Barros et al., Epithelia use butyrophilin-like molecules to shape organ-specific γδ T cell compartments. *Cell* **167**, 203–218.e17 (2016). doi: 10.1016/j.cell.2016.08.030; pmid: 27641500
63. M. Gomez de Agüero et al., The maternal microbiota drives early postnatal innate immune development. *Science* **351**, 1296–1302 (2016). doi: 10.1126/science.aad2571; pmid: 26989247
64. E. Leeansyah, L. Loh, D. F. Nixon, J. K. Sandberg, Acquisition of innate-like microbial reactivity in mucosal tissues during human fetal MAIT-cell development. *Nat. Commun.* **5**, 3143 (2014). doi: 10.1038/ncomms4143; pmid: 24452018
65. K. Hirota et al., Fate mapping of IL-17-producing T cells in inflammatory responses. *Nat. Immunol.* **12**, 255–263 (2011). doi: 10.1038/ni.1993; pmid: 21278737
66. A. Awasthi et al., Cutting edge: IL-23 receptor gfp reporter mice reveal distinct populations of IL-17-producing cells. *J. Immunol.* **182**, 5904–5908 (2009). doi: 10.4049/jimmunol.0900732; pmid: 19414740
67. L. E. Jao, S. R. Wente, W. Chen, Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13904–13909 (2013). doi: 10.1073/pnas.1308335110; pmid: 23918387
68. S. Conlan et al., *Staphylococcus* epidermidis pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. *Genome Biol.* **13**, R64 (2012). doi: 10.1186/gb-2012-13-7-r64; pmid: 22830599
69. I. R. Monk, I. M. Shah, M. Xu, M. W. Tan, T. J. Foster, Transforming the untransformable: Application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *mBio* **3**, e00277-11 (2012). doi: 10.1128/mBio.00277-11; pmid: 22434850
70. D. G. Gibson et al., Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009). doi: 10.1038/nmeth.1318; pmid: 19363495
71. D. E. Wright, A. J. Wagers, A. P. Gulati, F. L. Johnson, I. L. Weissman, Physiological migration of hematopoietic stem and progenitor cells. *Science* **294**, 1933–1936 (2001). doi: 10.1126/science.1064081; pmid: 11729320
72. M. Lochner, H. Wagner, M. Classen, I. Förster, Generation of neutralizing mouse anti-mouse IL-18 antibodies for inhibition of inflammatory responses in vivo. *J. Immunol. Methods* **259**, 149–157 (2002). doi: 10.1016/S0022-1759(01)00505-1; pmid: 11730850
73. J. A. Hall et al., Essential role for retinoic acid in the promotion of CD4(+) T cell effector responses via retinoic acid receptor α. *Immunity* **34**, 435–447 (2011). doi: 10.1016/j.immuni.2011.03.003; pmid: 21419664
74. B. J. Callahan et al., DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016). doi: 10.1038/nmeth.3869; pmid: 27214047
75. A. I. Dobin et al., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21 (2012). doi: 10.1093/bioinformatics/bts635; pmid: 23104886
76. S. Heinz et al., Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589 (2010). doi: 10.1016/j.molcel.2010.05.004; pmid: 20513432
77. S. Tripathi et al., Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. *Mol. Cell*, **18**, 723–735 (2015). doi: 10.1016/j.chom.2015.11.002; pmid: 26651948
78. M. Stoeckius et al., Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018). doi: 10.1186/s13059-018-1603-1; pmid: 30567574
79. T. Stuart et al., Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118

ACKNOWLEDGMENTS

We thank National Institute of Allergy and Infectious Diseases (NIAID) animal facility staff—in particular, D. Trageser-Cesler and C. Acevedo (NIAID Microbiome Program Gnotobiotic Animal Facility); J. Chen, N. Modi, and S. Mistry (NIH Center for Human Immunology); M. Smelkinson (NIAID Biological Imaging Section); T. Hawley and E. Stregevsky (NIAID Flow Cytometry Section); R. Awasthi (NCI Transgenic Mouse Model Laboratory); and G. Trinchieri, R. Salcedo, J. McCulloch, S. Sen, K. Beacht, J. LeGrand, E. Lewis, and J. Kehr. MR1 and CD1d tetramers were provided by the NIH Tetramer Core Facility. We acknowledge the NIAID Microbiome Program Sequencing Platform for performing the 16*S* rRNA and genomic sequencing. This study used the Office of Cyber Infrastructure and Computational Biology (OCICB) High Performance Computing (HPC) cluster at NIAID and the high-performance computational capabilities of the Biowulf Linux cluster at NIH. **Funding:** This research was supported in part by the Intramural Research Program of NIAID, NIH (ZIA-AI001115 and ZIA-AI001132 to Y.B.); the NIH Extramural Research Program (U19-AI111143 to J.A., R01-DK110174 and DP1-DK113598 to M.A.F., DP2-AI144968 to T.C.S.); the Cancer Research Institute Irvington Postdoctoral Fellowship Program (M.G.C. and I.V.-C.); the European Molecular Biology Organization (EMBO) Long-Term Fellowship Program and the ARC Foundation (S.T.); the National Institute of General Medical Sciences (NIGMS) Postdoctoral Research Associate Training (PRAT) Program (J.L.L.); Howard Hughes Medical Institute (HHMI) Hanna H. Gray Fellows Program (Y.E.C.); and the HHMI–Simons Faculty Scholars Program and the Chan Zuckerberg Biohub (M.A.F.). **Author contributions:** M.G.C. and Y.B. designed the study and experiments and wrote the manuscript; M.G.C. performed the experiments and analyzed the data; V.M.L., S.T., A.C.W., P.J.P.-C., S.-J.H., Y.E.C., K.L., S.F., A.W., S.R.K., I.V.-C., J.L.L., N.B., E.D.M., S.R., D.J.C., E.J.A., A.B., T.C.S., J.A., and M.A.F. participated in performing experiments, provided intellectual expertise, and/or helped to interpret experimental results; V.M.L. assisted with RNA-seq data analysis; S.T. and A.C.W. assisted with numerous experiments throughout; P.J.P.-C. isolated, cultured, and sequenced intestinal commensals; S.-J.H. performed confocal microscopy and parabiotic surgery; Y.E.C and M.A.F. generated the *S. epidermidis* NIHLM087 Δ*ribD* mutant; K.L. and J.A. synthesized 5-A-RU; S.F., A.W., and T.C.S. analyzed human skin biopsies; S.R.K. analyzed the presence of riboflavin synthesis genes; I.V.-C. analyzed the 16*S* rRNA data; J.L.L. assisted with wounding studies; N.B. coordinated GF experiments; E.D.M. participated in the conventionalization of GF mice; S.R. and E.J.A. generated MR1 tetramer used in initial experiments; D.J.C. provided antibody to IL-23R and isotype control; and A.B. coordinated the generation of *Mr1*$^{f/f}$ mice. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Reagents are available upon signing a materials transfer agreement. Antibody to IL-18 (SK113AE-4) is available from I. Förster under a materials transfer agreement with the University of Bonn. The RNA-seq data, 16*S* rRNA sequences, and bacterial genomes were deposited with the NCBI under accession nos. GSE128814 and PRJNA529261. All other data needed to evaluate the conclusions in this paper are present either in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/eaax6624/suppl/DC1
Figs. S1 to S5
Table S1

View/request a protocol for this paper from *Bio-protocol*.

11 April 2019; accepted 12 September 2019
10.1126/science.aax6624

RESEARCH ARTICLE SUMMARY

PLANT BIOLOGY

Dynamic control of plant water use using designed ABA receptor agonists

Aditya S. Vaidya*, Jonathan D. M. Helander*, Francis C. Peterson, Dezi Elzinga, Wim Dejonghe, Amita Kaundal, Sang-Youl Park, Zenan Xing, Ryousuke Mega, Jun Takeuchi, Bardia Khanderahoo, Steven Bishay, Brian F. Volkman, Yasushi Todoroki, Masanori Okamoto, Sean R. Cutler†

INTRODUCTION: Climate extremes create a need to mitigate the effects of drought on agriculture. The contributions of water to crop yield vary over a growing season but peak during reproductive development. Water banking strategies, which save soil water early in a growing season, reserve water for flowering and can increase yield under modest drought. Antitranspirants based on mimics of the phytohormone abscisic acid (ABA), which controls stomatal aperture, are sought as next-generation agrochemicals for controlling water use and increasing yield during drought.

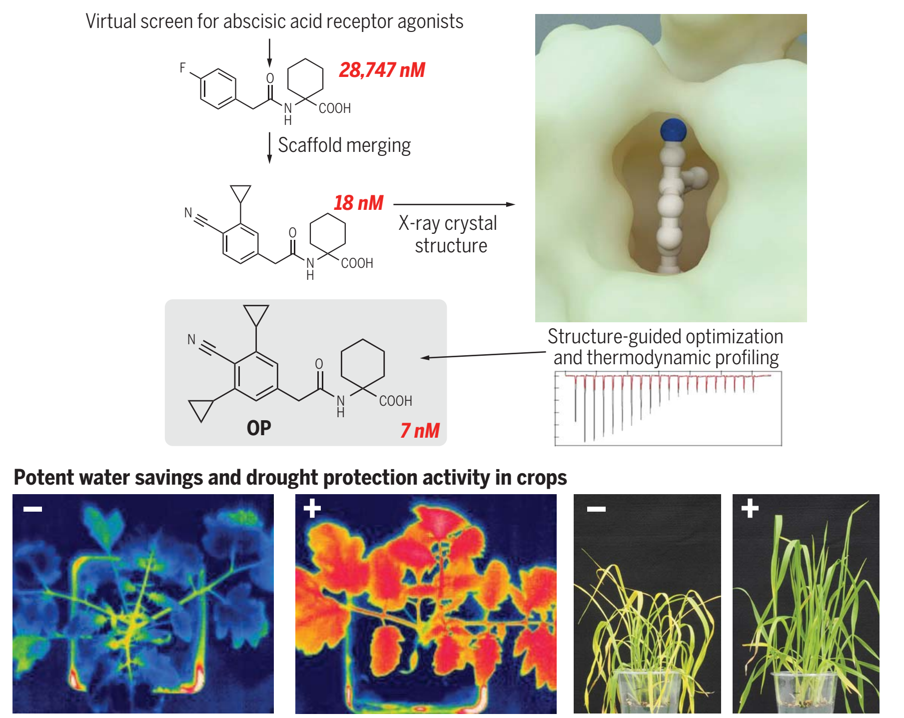
RATIONALE: Information on the structure and function of ABA receptors has created opportunities for agrochemical development. Current lead molecules have low and short-lived bioactivity in some relevant crop species, including wheat, the world's most widely grown staple crop. This liability is likely a consequence of incomplete activation of different ABA receptor subclasses. We reasoned that the idiosyncratic activity of these molecules was due, in part, to a lack of interaction between the agonist and a conserved lysine in ABA receptors that forms a salt bridge to ABA's carboxylate. We performed virtual screening to identify candidate agonists that interact with this lysine.

RESULTS: Two ABA receptor structures were used to screen against the ZINC database, a collection of commercially available ligands, using Glide docking protocols, requiring that hits interact with the conserved lysine. Candidate agonists were obtained and tested for receptor activation using in vitro assays. A family of substituted phenyl acetamido-cyclohexane carboxylic acid ABA receptor panagonists was identified. Scaffold merging was used to improve binding: We grafted an optimized headgroup of an existing sulfonamide onto our phenyl acetamido-cyclohexane carboxylic acid scaffold to yield a chimeric ligand (3CB) that displayed an improvement toward target sites of up to three orders of magnitude. Analysis of a 3CB-PYL10 crystal structure suggested that addition of appropriately situated hydrophobes to 3CB might improve activity. A 3CB derivative was synthesized, yielding an agonist that we have named opabactin (OP) for overpowered ABA receptor activation. Thermodynamic characterization of 3CB or OP receptor binding reactions indicates that the newly generated scaffold's improvements are enthalpically driven relative to sulfonamides, consistent with the salt bridge observed in our crystal structure. Biological studies show that OP is ~10-fold more active in inhibiting seed germination (a response driven by ABA) than ABA itself. Experiments in wheat, tomato, barley, *Arabidopsis*, and *Commelina* demonstrate bioactivity in vegetative tissues across diverse species. Time course experiments in wheat and tomato using thermal imaging show that OP induces a more sustained antitranspirant response than ABA, and they document poor activity of sulfonamide agonists in those two species. To understand which receptors are necessary for OP action, we used *Arabidopsis* mutant strains to show that OP requires the subfamily III receptors PYR1, PYL1, and PYL2 for maximal activity. Thus, our virtual screening experiments yielded an ABA receptor agonist that functions as an antitranspirant.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaw8848

CONCLUSION: A newly generated ABA agonist scaffold was identified and optimized through a structure-guided approach. The chemical biology of ABA receptor agonists can be broadened by designing ligands that engage a conserved lysine residue in the binding pocket. This pharmacophoric feature results in a favorable enthalpic binding profile and lower dissociation constants than existing sulfonamide-based ligands. Our ABA agonist, opabactin, has activity in multiple monocots and eudicots and addresses the limitations of existing sulfonamide molecules being explored as tools for mitigating the effects of drought on crop yields. ■



Water savings and drought protection activity in crops. Virtual screening yielded a scaffold 3B4 whose potency was improved by ~1600-fold by scaffold merging, yielding 3CB. Structure-based optimization gave opabactin (OP). Thermodynamic profiling shows that OP binding to its targets PYR1 and HAB1 is enthalpically driven with about one-tenth the dissociation constant K_d of ABA. OP exhibits more potent and longer-lasting antitranspirant effects than existing sulfonamide-based ligands.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: sean.cutler@ucr.edu
Cite this article as A. S. Vaidya *et al.*, *Science* **366**, eaaw8848 (2019). DOI: 10.1126/science.aaw8848

RESEARCH ARTICLE

PLANT BIOLOGY

Dynamic control of plant water use using designed ABA receptor agonists

Aditya S. Vaidya[1,2]*, Jonathan D. M. Helander[1,2]*, Francis C. Peterson[3], Dezi Elzinga[1,2], Wim Dejonghe[1,2], Amita Kaundal[1,4], Sang-Youl Park[1,2], Zenan Xing[1,2], Ryousuke Mega[5], Jun Takeuchi[6,7], Bardia Khanderahoo[1,2], Steven Bishay[1,2], Brian F. Volkman[3], Yasushi Todoroki[6,7], Masanori Okamoto[8,9], Sean R. Cutler[1,2]†

Drought causes crop losses worldwide, and its impact is expected to increase as the world warms. This has motivated the development of small-molecule tools for mitigating the effects of drought on agriculture. We show here that current leads are limited by poor bioactivity in wheat, a widely grown staple crop, and in tomato. To address this limitation, we combined virtual screening, x-ray crystallography, and structure-guided design to develop opabactin (OP), an abscisic acid (ABA) mimic with up to an approximately sevenfold increase in receptor affinity relative to ABA and up to 10-fold greater activity in vivo. Studies in *Arabidopsis thaliana* reveal a role of the type III receptor *PYRABACTIN RESISTANCE-LIKE 2* for the antitranspirant efficacy of OP. Thus, virtual screening and structure-guided optimization yielded newly discovered agonists for manipulating crop abiotic stress tolerance and water use.

Water flux from soil to atmosphere occurs primarily by its movement through plants (*1*). Water enters the plant body through roots and exits as vapor from leaf surface stomata, small pores formed by neighboring guard cells that open and close to regulate gas exchange. The accumulation of biomass by photosynthesis requires CO_2 to access inner leaf mesophyll cells through stomata, but this comes at the cost of H_2O loss, which is driven by the difference in water vapor pressure between the inner leaf and atmosphere. Plants, therefore, face a trade-off between water use and growth. One consequence of this trade-off is that attempts to improve water use efficiency often come at the cost of reduced growth and yield (*2*). Conversely, studies of high-yielding wheat varieties have shown that they transpire more water than lower-yielding ones, and reduced water use efficiency has been proposed to be a consequence of selection for high grain yields in wheat (*3*, *4*).

It is, therefore, intrinsically challenging to create crops with both high yield and high water use efficiency. Small molecules can, in principle, address this dilemma by coaxing high-yielding plant varieties with low water use efficiency into more water-efficient states on demand (*5*). This strategy would also enable data streams from precision agriculture to be directed toward conserving limited water resources and responding to emergent adverse events. Moreover, given that ~11% of Earth's surface is dedicated to crop production (*6*) (mostly nonirrigated), better tools for managing crop water use could have global effects on water usage and help to address the ~$29 billion U.S. dollars in agricultural losses attributed to drought annually (*7*).

The scope and global importance of water use efficiency is driving the development of next-generation agrochemicals for managing responses to drought (*8*–*20*). These efforts have focused primarily on small-molecule mimics of the phytohormone abscisic acid (ABA), which coordinates plant physiology, growth, and stress responses with water availability. ABA responses are mediated by a signaling module that involves soluble *Pyrabactin Resistance 1/PYR1-Like/Regulatory Component of ABA Receptor* (PYL) ABA receptors, clade A type 2C protein phosphatases (PP2Cs), and subfamily three SNF1-related kinases (SnRK2s) (*21*, *22*). When ABA binds to its soluble receptors, PYLs, a conformational change enables the receptors to bind to PP2C active sites and inhibit phosphatase activity. This, in turn, leads to the accumulation of activated SnRK2 kinases, which phosphorylate downstream factors and trigger physiological responses such as guard cell closure. This system is thus a target for both genetic and agrochemical control of drought tolerance. Overexpression of ABA receptors can increase water use and photosynthetic efficiency in both *Arabidopsis* and wheat with negligible effects on growth in controlled environment studies (*23*, *24*), such genetic interventions can complement chemical strategies because the two can be combined for improved effects (*25*).

The most potent synthetic ABA receptor agonists to date are ABA analogs or molecules that share a core sulfonamide linkage related to quinabactin (*18*, *20*, *26*–*30*). Flowering plant ABA receptors cluster into three phylogenetically distinct subfamilies (*31*, *32*). Genetic analyses in *Arabidopsis* have demonstrated that quinabactin and other sulfonamide agonists require subfamily IIIA receptors (PYR1 and PYL1) to exert their antitranspirant effects (*26*, *33*), but it is unclear whether activation of other ABA receptors, which in many angiosperms exceed a dozen per genome, would enhance antitranspirant effects. Although quinabactin activates PYR1 with nanomolar potency, it shows reduced or negligible activity on other receptors (*26*), including PYL4, one of the key receptors expressed in *Arabidopsis* guard cell microarray experiments (*34*, *35*), a liability that could potentially limit efficacy. Upon investigating the effects of quinabactin and other sulfonamides in wheat and tomato, we observed low and short-lived bioactivity, which we speculated was the result of weak affinity for necessary targets. A structure-guided ligand discovery and optimization campaign led us to opabactin (OP), a potent agonist with activity in wheat and tomato. Genetic analyses show that PYR1, PYL1, and PYL2 are key target sites necessary for OP's potent antitranspirant activity.

Discovery of a nonsulfonamide ABA pan agonist scaffold by virtual screening

In ABA/receptor cocrystal structures, ABA's carboxylate forms a salt bridge with a conserved lysine (K59 in PYR1) that current sulfonamide agonists are unable to form (*33*, *36*–*39*) (Fig. 1A). We reasoned that the idiosyncratic activation of different ABA receptor types was due, in part, to lack of K59 engagement. We therefore used Glide docking protocols (*40*) to screen a collection of available ligands (*41*) for agonists that interact with K59. A subset of the predicted binders were tested for activity in vitro, using a pool of diverse recombinant ABA receptors (*Arabidopsis* PYL1, PYL2, PYL4, and PYL8, and maize PYL8). This process yielded 22 hits (half maximal inhibitory concentration, $IC_{50} \leq 25$ μM) (Fig. 1B, fig. S1, and table S1), most of which could be clustered into three scaffolds: a group of sulfonamides reminiscent of pyrabactin and quinabactin, a family of substituted aminopropanediols, and a set of amino acid amide conjugates (Fig. 1B, fig. S1, and table S1). We focused optimization efforts on the amide scaffold because of its

[1]Institute for Integrative Genome Biology, University of California, Riverside, Riverside, CA 92521, USA. [2]Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA 92521, USA. [3]Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226, USA. [4]Department of Plants, Soils and Climate, Utah State University, Logan, UT 84322, USA. [5]Arid Land Research Center, Tottori University, 1390 Hamasaka, Tottori 680-0001, Japan. [6]Faculty of Agriculture, Shizuoka University, Shizuoka 422-8529, Japan. [7]Research Institute of Green Science and Technology, Shizuoka University, Shizuoka 422-8529, Japan. [8]Center for Bioscience Research and Education, Utsunomiya University, 350 Mine, Utsunomiya, Tochigi 321-8505, Japan. [9]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.
*These authors contributed equally to this work.
†Corresponding author. Email: sean.cutler@ucr.edu

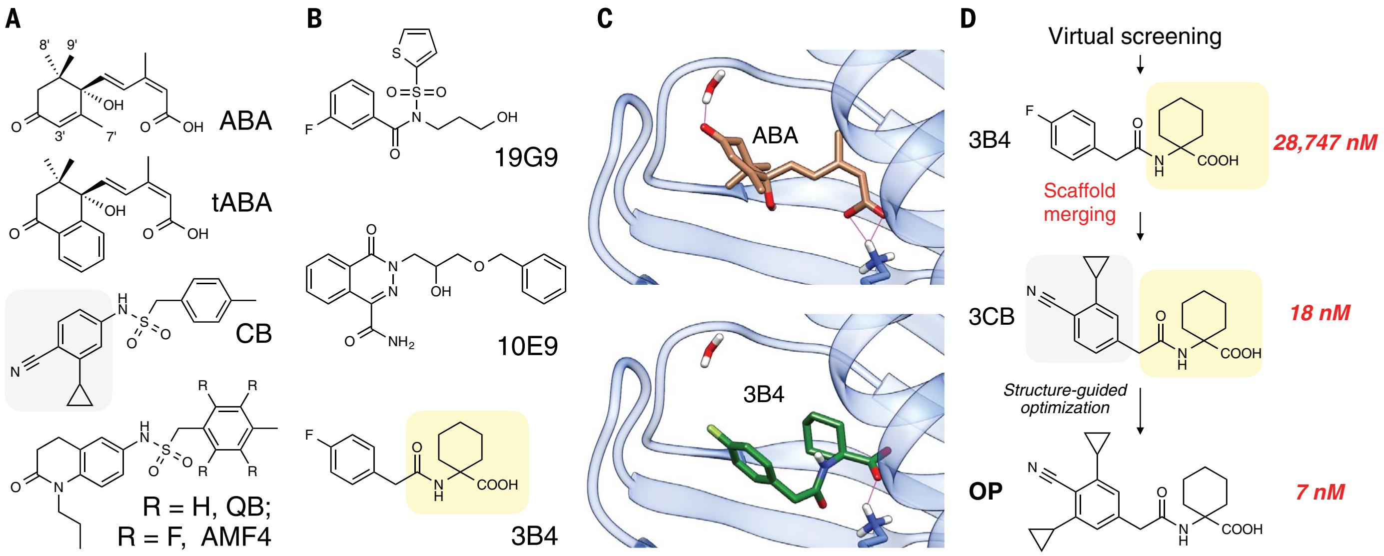


Fig. 1. Virtual screening and scaffold merging enabled the design of overpowered ABA receptor agonists. (**A**) Structures of existing ABA receptor agonists; the gray box highlights the 4-cyano-3-cyclopropylphenyl head group used in our scaffold merging strategy. tABA, tetralone ABA; CB, cyanabactin; QB, quinabactin; R, radical. (**B**) Structures of representative hits (IC_{50} ≤ 25 μM) belonging to the different scaffolds identified; the yellow box highlights the amide scaffold that was merged with the cyanabactin headgroup. The complete list of hits identified is provided in fig. S1 and table S1. (**C**) Binding modes of ABA and hit 3B4, as originally predicted by docking to PYR1; the figure illustrates that 3B4 lacks a hydrogen bond acceptor for the Trp-lock water with thin red lines representing hydrogen bonds from 2.7 to 3.1 Å. (**D**) Hit discovery and optimization. The ZINC collection of 18 million purchasable drug-like compounds was screened against PYR1 and PYL9 to identify candidate binders that interact with K59. Of the top 10,000 hits, 1700 were obtained and tested for activity in vitro. The best hit obtained, 3B4, lacked an H-bond acceptor for the Trp lock. We therefore installed the cyanabactin head group on the 3B4 amide scaffold. A subsequent structure of 3CB-PYL10 suggested that bioactivity could be improved through the addition of an additional hydrophobe to interact with the receptors' 3′ tunnels, which yielded the agonist opabactin (OP). IC_{50} values shown on the right (in red) are for the compounds tested against subfamily III receptor PYL2 (see Fig. 2 and fig. S2 for complete IC_{50} data). Synthetic schemes for the synthesis of 3CB and OP are provided in fig. S2.

relatively high activity across the eleven *Arabidopsis* receptors tested (tables S1 and S2).

Scaffold merging yields a potent panreceptor agonist

The most active amide identified [the chemical in screening plate 3, well B4 (3B4)] displayed submicromolar IC_{50} values and apparent panreceptor agonism (table S2). Our virtual screening data predicted that its carboxylate could form a salt bridge to K59 but that it lacked a hydrogen bond acceptor needed to engage the tryptophan (Trp)–lock water, which interacts with ABA's ring ketone and coordinates an H-bond network that stabilizes activated receptors (Fig. 1C) (*42*–*44*). In previous work, we showed that this function could be served by a nitrile in the designed ligand cyanabactin (Fig. 1A) (*33*). We hypothesized that a chimeric ligand possessing cyanabactin's 4-cyano-3-cyclopropylphenyl headgroup grafted onto the 3B4 amide scaffold would improve binding affinity by providing simultaneous interactions with K59 and the Trp lock, essentially adopting a "scaffold merging" approach that has been exploited in the design of heat shock protein 90 inhibitors (*45*) and strigolactone mimics (*46*). To test this, we synthesized and characterized the chimeric 3B4-cyanabactin hybrid compound 3CB (scheme shown in fig. S2 and the materials and methods). We also synthesized AMF4 (ABA mimic 1 tetrafluoro derivative), a quinabactin derivative and sulfonamide agonist with improved potency (*25*) and tetralone ABA (*29*, *30*), an ABA derivative with increased potency, so that we could compare these other potent agonists to our chimeric ligand.

Our synthesized chimeric ligand (3CB) displayed an increase of up to three orders of magnitude in affinity for subfamily III receptors and low nanomolar pan agonist activity (Fig. 2A, fig. S3A, and table S2), validating the scaffold-merging strategy used. To establish whether the agonist was active in monocots, which diverged from the eudicot lineage ~200 million years ago and include the world's major grain crops, we tested the sensitivity of wheat ABA receptors to 3CB using a panel of subgenome D ABA receptors and a native wheat PP2C (*24*). These data show that 3CB is a similarly potent pan agonist of wheat receptors, unlike sulfonamide agonists (Fig. 2B, fig. S3B, and table S3). Moreover, we observe that the chimeric ligand is approximately twofold more active than ABA in *Arabidopsis* seed germination assays, indicating adequate bioavailability (Fig. 2C and fig. S4). Collectively, these experiments demonstrate that the combined use of virtual screening and scaffold merging enabled us to discover and optimize an agonist that binds multiple ABA receptors and is active in both eudicots and monocots.

Structure-guided optimization of an overpowered agonist

To understand the atomic basis of 3CB's high activity, we cocrystallized it with the *Arabidopsis* subfamily I receptor PYL10, which was chosen because it expresses in *Escherichia coli* at high levels relative to other subfamily I receptors, and solved an x-ray crystal structure of a $PYL10^{25-183}$:3CB complex using molecular replacement at a resolution of 2.4 Å (table S4). The $PYL10^{25-183}$:3CB complex crystallized in the space group $P3_121$ and contained a single protomer in the asymmetric unit. Several rounds of structural refinement were carried out before modeling 3CB into the ligand-binding pocket's unbiased electron density (fig. S5). A real-space correlation coefficient of 0.95 calculated between the unbiased electron density and 3CB indicates agreement between the model and observed electron density. These data position 3CB in the PYL10 ligand-binding pocket with its carboxylate forming a salt bridge to K56 (homologous to PYR1's K59), and its aryl ring oriented toward the gate and latch loops, mimicking ABA's cyclohexenone ring and gem-dimethyl group (Fig. 3A). Direct

A

		ABA	QB	AMF4	tABA	3CB	OP
III	PYR1	149	140	47	122	48	13
	PYL1	160	152	47	115	17	15
	PYL2	36	352	144	40	18	7
	PYL3	12	1689	504	14	14	7
II	PYL4	27	>50000	24394	22	56	74
	PYL5	7	301	119	10	6	8
	PYL6	6	>50000	>50000	6	20	32
	PYL11	67	–	>50000	26	138	98
I	PYL8	57	–	–	98	13	–
	PYL9	43	–	–	31	49	–
	PYL10	148	–	–	169	18	–

Arabidopsis: receptor-mediated HAB1 inhibition IC_{50} (nM)

B

		ABA	QB	AMF4	tABA	3CB	OP
III	TaPYL1	128	196	75	94	23	7
	TaPYL2	76	1151	273	52	45	8
	TaPYL3	117	3941	620	114	84	15
II	TaPYL4	26	>10000	2102	15	12	16
	TaPYL5	9	8901	2686	9	11	13
	TaPYL6	12	4553	294	7	7	6
	TaPYL7	42	>10000	4795	32	20	18
I	TaPYL8	227	–	–	58	597	94

Triticum: receptor / TaPP2C IC_{50} (nM)

C

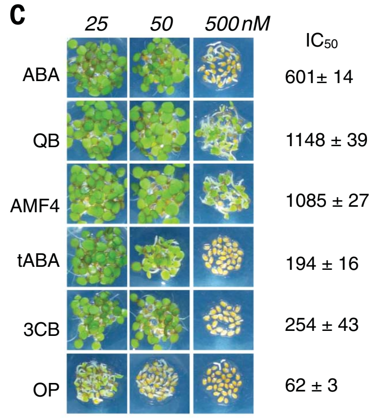


Fig. 2. 3CB and OP are best-in-class ABA receptor agonists. (**A**) Agonist potency against different *Arabidopsis* receptors, as measured using agonist/receptor-mediated inhibition of ΔN-HAB1 phosphatase activity (n = 3 replicates). A dash indicates that PP2C activity was >75% control at 50 μM test chemical (the highest concentration tested). The colored boxes indicate the subfamily membership of the receptors assayed. (**B**) Agonist potency against different wheat receptors, using indicated receptors and TaPP2C. A dash indicates that PP2C activity was >75% control at 10 μM test chemical (the highest concentration tested). We note that TaPYL9 is a pseudoreceptor with an atypical binding pocket. (**C**) OP shows an ~10-fold increase in in vivo potency relative to ABA. The potency of different agonists on *Arabidopsis* seed germination and their corresponding IC_{50} values (concentrations required to inhibit germination by 50%). Photographs were taken 4 days post stratification (n = 6 replicates). The full-dose response curves used to infer IC_{50} values for all three figure panels are shown in figs. S3 and S4 and tables S2 and S3.

polar contact between N163's amide NH and 3CB's amide carbonyl, and two water-mediated contacts to E90 and F98 (Fig. 3A), additionally stabilize 3CB-PYL10 interactions. 3CB's cyclohexyl ring packs against L159, V156, Y116, and I106, which line the C6 cleft, a small hydrophobic pocket that normally interacts with ABA's C6 methyl and is a ligand-binding hotspot (*47*).

On the basis of our prior crystallographic analyses of a PYR1-cyanabactin complex (*33*), we anticipated that 3CB's cyclopropyl substituent might extend into the hydrophobic 3′ tunnel, which contacts ABA's C7′ methyl (Fig. 3B). Instead, our structure shows 3CB's cyclopropyl group oriented away from the 3′ tunnel, packing against the gate and latch loops (Fig. 3A). The vacancy of the 3′ tunnel, therefore, suggested that we might increase 3CB's potency by adding hydrophobic substituents to interact with this pocket, a strategy used to improve other agonists (*33*, *48*). To test this idea, we synthesized a dicyclopropyl 3CB analog, which we have named opabactin (OP) for overpowered ABA receptor activation (synthetic details are provided in fig. S2). We tested OP activity in vitro using a receptor-agonist N-terminal deletion of HAB1 (ΔN-HAB1) inhibition assay and observed that it activates *Arabidopsis* subfamily III receptors with potencies that are up to ~10-fold greater than ABA's; in vivo, OP is ~10-fold more bioactive, as measured using *Arabidopsis* seed germination assays (Fig. 2C), and induces an approximately fivefold stronger effect than ABA on guard cells, as measured using a *Commelina* guard cell closure assay (fig. S6). Although activity against subfamily III receptors improved, activity against *Arabidopsis* subfamily I receptors diminished (Fig. 2A). In contrast, we observed that OP activates the wheat subfamily I receptor TaPYL8 and is thus a pan agonist of the wheat ABA receptors tested (Fig. 2B). Inspection of sequence alignments reveals a single amino acid difference between *Arabidopsis* and wheat subfamily I ligand-binding pockets (a bulky leucine is replaced by a smaller valine in TaPYL8); moreover, this residue packs tightly against the location of the second cyclopropyl in OP and would cause steric clash in *Arabidopsis* receptors (Fig. 3, C and D). Yeast two-hybrid assays confirm that the leucine-to-valine pocket difference determines subfamily I sensitivity to OP (Fig. 3E). Thus, the 3CB-PYL10 cocrystal structure enabled us to increase its activity against subfamily III targets and yielded an agonist that is ~10-fold more potent than ABA in seeds.

OP-PYR1 binding is enthalpically driven

The 3CB carboxylate-K56 salt bridge is likely to improve binding affinity via an enthalpic contribution. To examine this, we conducted isothermal titration calorimetry (ITC) experiments comparing 3CB- and OP-PYR1 binding isotherms (Fig. 4A). Prior analyses (*47*) revealed an apparent dissociation constant (K_d) of 4.8 μM for quinabactin-PYR1 binding and indicated that quinabactin binding is entropy driven (change in enthalpy, $\Delta H = -2.8$; change in entropy, $-T\Delta S = -4.7$ kcal/mol) (Fig. 4A). In contrast, 3CB- and OP-PYR1 binding reactions are more enthalpic and of higher affinity relative to quinabactin and cyanabactin (OP K_d = 1.9 μM; quinabactin K_d = 4.8 μM) (Fig. 4A). The presence of the second cyclopropyl substituent in OP relative to 3CB provides an additional −1.7 kcal/mol ΔH for PYR1 binding; this may indicate protein movement in response to ligand binding and/or be a consequence of van der Waals contacts, which in some systems are sufficient to create exothermic binding isotherms without conformational changes in protein structure (*49*–*51*). We additionally examined ligand–PYR1-HAB1 binding reactions, comparing ABA, quinabactin, and OP (Fig. 4B). These experiments show that OP's K_d for PYR1-HAB1 binding is approximately one-seventh that of ABA (28 ± 5 nM versus 201 ± 14 nM). This increase in binding affinity translates into improved potency of OP in *Arabidopsis* seed germination assays described above (Fig. 2C). Collectively, our ITC data show that our newly identified amide scaffold provides more enthalpy-driven, higher-affinity binding than previous sulfonamides, a feature of many best-in-class ligands (*52*, *53*).

OP has antitranspirant activity and activates ABA signaling in vegetative tissues

Our initial tests indicated that both 3CB and OP activate multiple ABA responses in *Arabidopsis* vegetative tissues (fig. S7) and demonstrate antitranspirant activity in the grain crop barley (fig. S8). To investigate OP's effects in other crop species, we treated tomato and wheat plants with different concentrations of ABA, quinabactin, AMF4, or OP and used thermal imaging to measure their effects on transpiration for up to 5 days after application (relative to mock-treated controls). In both tomato and wheat, ABA's effects on leaf temperature dissipate after 2 to 3 days, and sulfonamide agonist effects are no longer evident after 48 hours (Fig. 5). In contrast, OP's effects

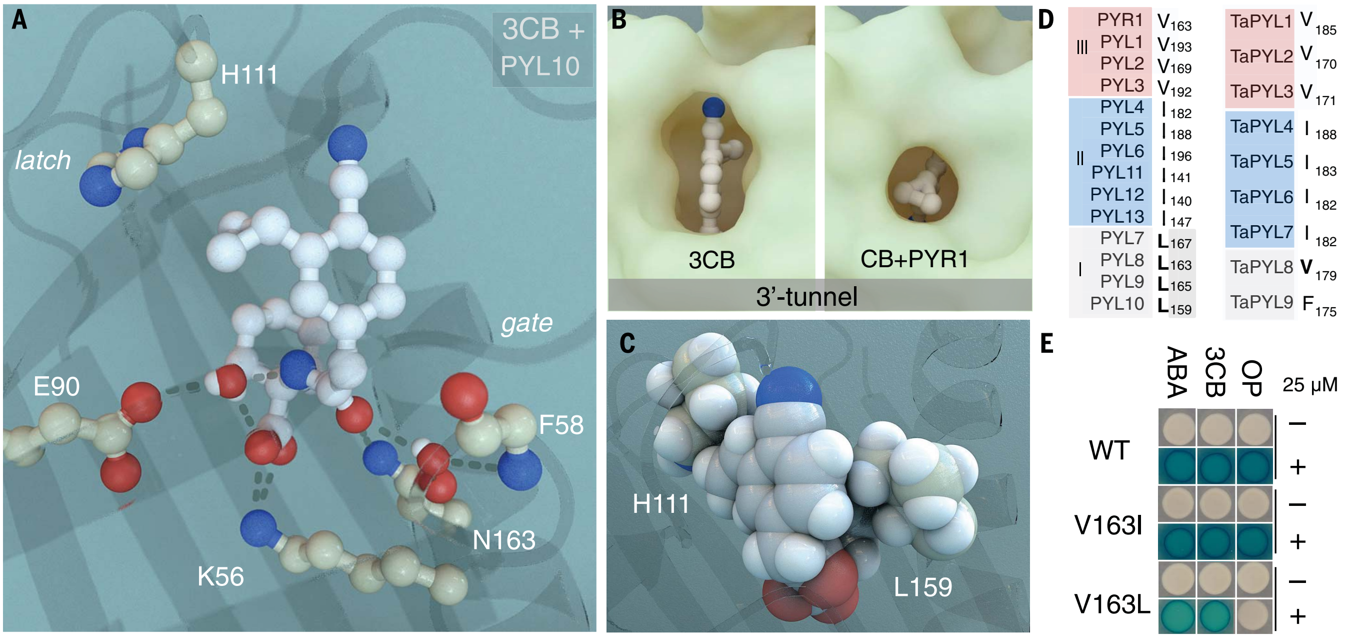
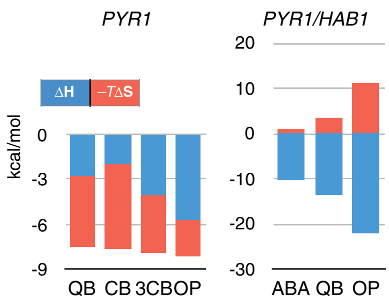


Fig. 3. Structure of 3CB bound to PYL10. (**A**) Cocrystal structure of 3CB bound to PYL10, dashes representing hydrogen bonds from 2.6 to 3.3 Å. (**B**) Differences in orientation of the cyclopropyl group present in the 3CB:PYL10 structure versus the CB:PYR1 structure (PDB ID 5UR5). (**C**) Positioning of H111 and L159 in PYL10 relative to 3CB's 3-cyclopropyl-4-cyanobenzyl headgroup. This illustrates the site of predicted steric clash that would occur between L159 and OP's C5 cyclopropyl substituent when bound to *Arabidopsis* subfamily I receptors. (**D**) Sequence variation at the selectivity-determining residue (homologous to PYL10 L159), which is the only residue lining the 3′ tunnel that is variable between ABA receptors). (**E**) Genetic evidence that sequence variation at the selectivity-determining residue (L159 in PYL10) determines OP selectivity. V163 in PYR1 is homologous to L159 in PYL10. To test the role of this residue in OP selectivity, wild-type PYR1, V163I, and V163L were tested for interactions with HAB1 using an established yeast two-hybrid assay (chemicals tested at 25 μM); the bulkier leucine substitution corresponding to that present in *Arabidopsis* subfamily I receptors blocks OP action. Single-letter abbreviations for the amino acid residues are as follows: E, Glu; F, Phe; H, His; I, Ile; K, Lys; L, Leu; N, Asn; V, Val.

A

Ligand	Receptor	ΔH (kcal/mol)	−TΔS	K_d (nM)	
QB	PYR1	-2.8 ± 0.1	-4.7 ± 0.2	4780 ± 1000	Ref. 47
CB	PYR1	-2.0 ± 0.3	-5.6 ± 0.3	3400 ± 500	Ref. 33
3CB	PYR1	-4.0 ± 0.1	-3.9 ± 0.1	2630 ± 590	This study
OP	PYR1	-5.7 ± 0.1	-2.4 ± 0.2	1850 ± 260	This study
ABA	PYR1+HAB1	-10.1 ± 0.3	0.7 ± 0.3	201 ± 14	This study
QB	PYR1+HAB1	-13.5 ± 0.6	3.6 ± 0.6	103 ± 8	This study
OP	PYR1+HAB1	-21.9 ± 2.2	11.2 ± 2.3	28 ± 5	This study



B

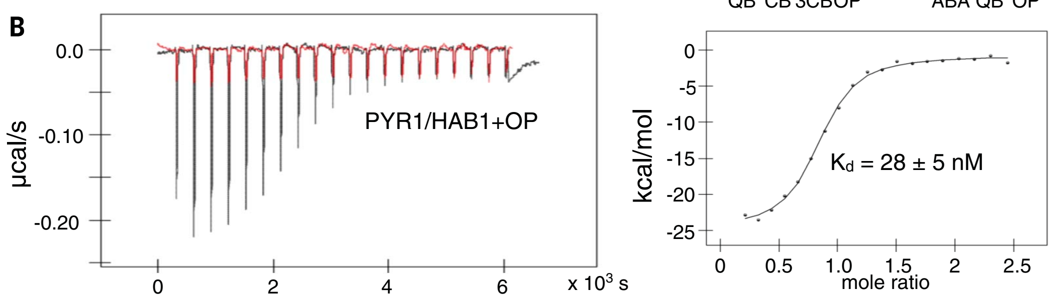


Fig. 4. The high binding affinities of 3CB and OP are enthalpy driven, in contrast to the entropy-driven binding of sulfonamides. (**A**) Measured thermodynamic parameters of binding of agonists to PYR1 in the presence or absence of HAB1; the top two rows are previously published observations, and all the rest were determined in this study. The bar graphs summarize the ITC data and illustrate that sulfonamide agonist binding is driven primarily by entropy, whereas the 3CB and OP binding are more enthalpically driven. (**B**) Representative thermogram and binding isotherm. ITC experiments were conducted by repeated injections of 2.5 μl of OP (60 μM) into a 1:1 mixture of PYR1:ΔN-HAB1 (10 μM). Binding isotherms were fitted to an independent one-site binding model using NanoAnalyze software (TA Instruments, USA); n = 3 replicate experiments, standard deviations shown. The red trace is a control of buffer injections (rather than buffer + ligand) into PYR1–ΔN-HAB1. The data are consistent with high-affinity binding of OP to the receptor complex; however, precise K_d values in the low nanomolar range are challenging to deduce.

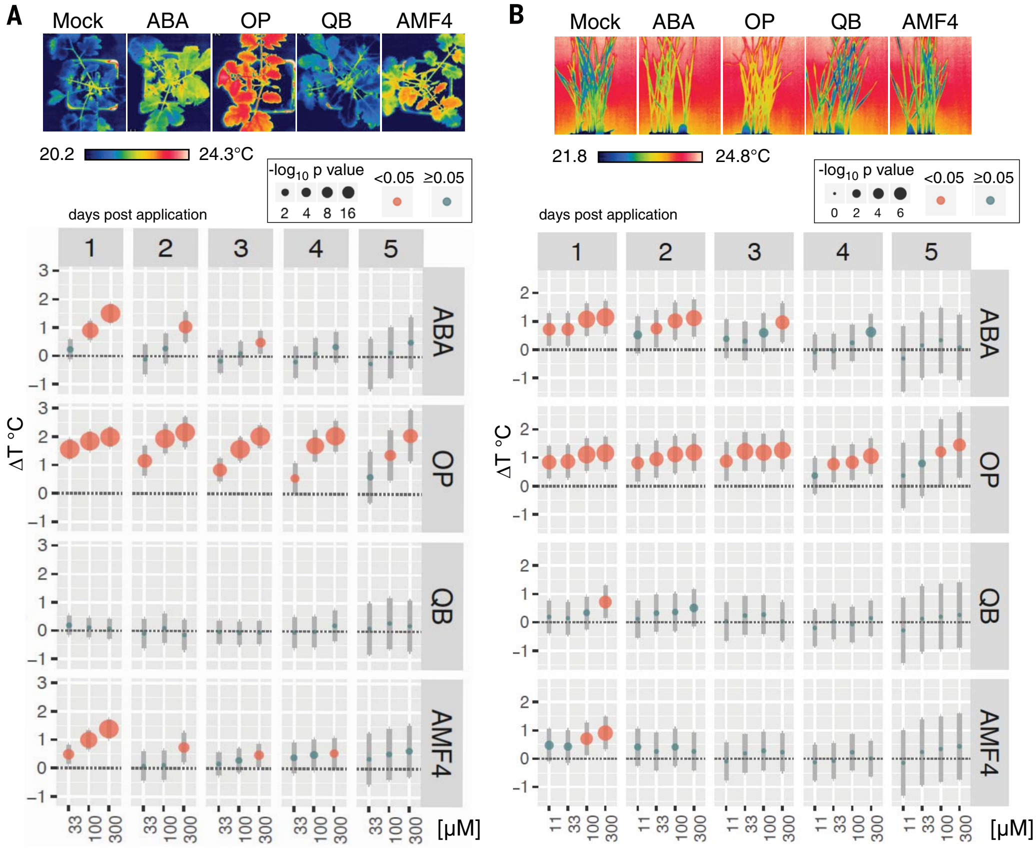


Fig. 5. OP has bioactivity and long-lasting effects in tomato and wheat, whereas sulfonamide agonists do not. (**A**) Infrared images of 6-week-old tomato seedlings (UC82) treated with 100 μM of different chemical solutions composed of 0.5% dimethyl sulfoxide (DMSO) and 0.05% Silwet-77. Graphs depict the variation in difference in temperature levels of treated plants relative to mock-treated plants up to 5 days after application (n = 18 replicates for mocks and n = 6 for chemical treatment). (**B**) Infrared images of 4-week-old wheat seedlings (WB-9229) sprayed with 11 μM of different chemicals solution composed of 0.5% DMSO and 0.05% Silwet-77. Graphs depict the variation in difference in temperature levels of treated plants, relative to mock-treated plants up to 5 days after application. Statistical analyses were conducted in R using a one-way analysis of variance (ANOVA) and post-hoc Dunnett tests to obtain multiplicity adjusted P values for treatment effects relative to mock-treated controls (n = 10 replicates for mocks and n = 5 for chemical treatment). 95% confidence intervals are shown.

remained evident at 5 days (the latest time point for which we collected data) in both species (Fig. 5). This increased bioactivity is likely the result of a combination of improved persistence and target site potency; however, metabolic studies will be required to resolve this. Whatever the case, these data show that OP has longer-lasting effects relative to ABA and other agonists in wheat, a globally important staple crop.

We next examined OP's effects on both ABA-induced transcriptional responses and drought tolerance in wheat. Using a progressive water stress regime, we observed that wheat seedlings treated with ABA, 3CB, and OP delayed wilting under water stress and retained more chlorophyll than did mock-treated plants, indicating less stress-induced senescence in treated plants relative to mock-treated controls (Fig. 6, A and B). In contrast, quinabactin-treated plants were only modestly protected under the same conditions, consistent with the weak activity of sulfonamide agonists described above (Fig. 6) and results from direct receptor-based assays (Fig. 2B). Quantitative reverse transcription polymerase chain reaction (qRT-PCR) experiments of treated seedlings for three ABA-responsive genes also demonstrate that OP shows increased activity relative to ABA and quinabactin (Fig. 6, C to E). Collectively, these data show that OP's high binding affinity for ABA receptors (relative to ABA and sulfonamides) translates into bioactivities that exceed both ABA and sulfonamide agonists in planta.

Subfamily III receptors are crucial for mediating the effects of OP in planta

To examine the receptor dependence of OP's action, we examined agonist-induced responses in wild-type or mutant strains lacking the major members of the different receptor subtypes. The effects of quinabactin and cyanabactin on *Arabidopsis* transpiration, germination, and gene expression are mediated primarily by the subfamily III receptors PYR1 and PYL1, which form a subclade within family III (*33*). To examine if this subclade similarly executes OP's effects, we examined the sensitivity of wild type; *pyr1,pyl1* (referred to as *0;1*); *pyr1, pyl1,pyl2* (*0;1;2*, subfamily III); *pyl4,pyl5* (*4;5*, subfamily II); and *pyl8,pyl9* (*8;9*, subfamily I) mutant strains to OP and other agonists. Both 3CB and OP inhibit *Arabidopsis* seed germination and seedling establishment, but the effects of OP are largely absent in the *0;1;2* triple mutant strain, while those of 3CB are not (Fig. 7A). These genetic observations therefore suggest that OP's action on seeds is mediated by subfamily III receptors because OP-treated seeds germinate in the *0;1;2* mutant. As expected, quinabactin and AMF4 effects on seeds are absent in the *0;1* double mutant; however, OP retains bioactivity in the *0;1* mutant, and its effects are only greatly reduced in the *0;1;2* triple mutant (Fig. 7A). This demonstrates that, in contrast to previously characterized sulfonamide agonists, PYL2 is additionally necessary for full OP action in seeds. Collectively these results further confirm the importance of subfamily III receptors for sulfonamide agonist action and illuminate an added role for PYL2 in OP's action.

To determine whether the involvement of PYL2 in OP's action also applies to vegetative responses, we examined OP's effects on leaf temperature in wild type, *0;1*, *0;1;2*, and *pyl2* mutant seedlings. These experiments show that OP's activity in leaf tissues can be greatly reduced in the *0;1;2* triple mutant but not the *0;1* double mutant strain, paralleling our observations in seeds (Fig. 7, B and C). Furthermore, both ABA and OP inhibit primary root growth and greening in *Arabidopsis* seedlings in a dose-dependent manner, as expected, and OP's effects in these growth assays can be greatly reduced in the *0;1;2* triple mutant (fig. S9). These genetic analyses demonstrate that PYL2, in addition to PYR1 and PYL1, is a cellular target that can be activated by synthetic agonists to manipulate plant transpiration and ABA responses.

Discussion

Synthetic ABA mimics are being investigated as tools for mitigating the impact of drought

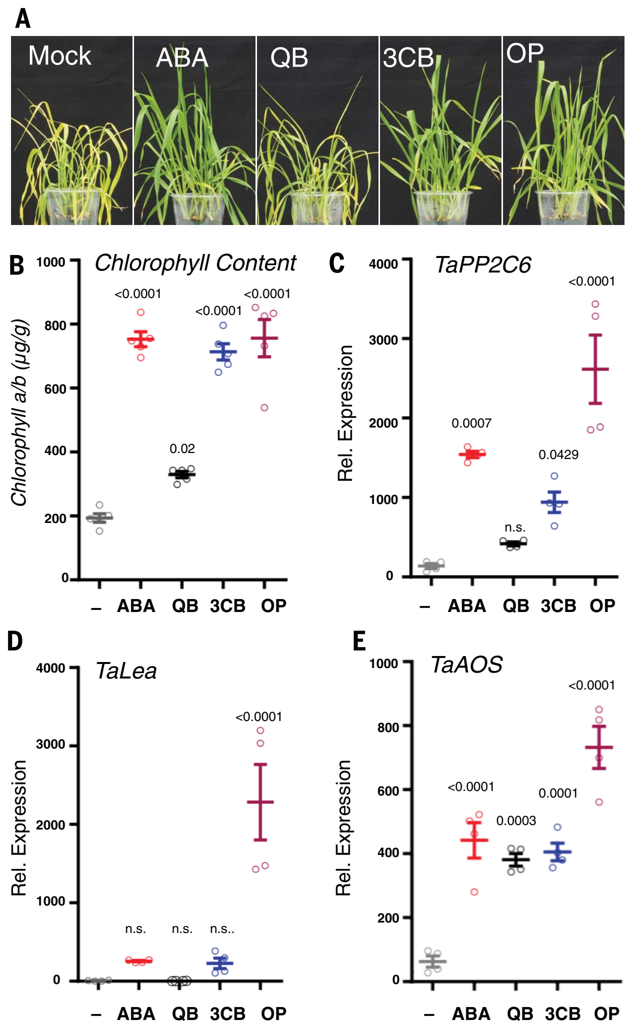


Fig. 6. OP enhances drought tolerance and hyperactivates ABA transcriptional responses in wheat. (**A**) Effects of ABA receptor agonists in wheat. Two-week-old plants were sprayed with chemical solution containing 0.05% Tween-20 and 50 µM compounds. After 3 days of water deprivation, plants were rewatered, and stress recovery images were acquired after 12 hours. (**B**) Chlorophyll content of aerial parts of plants after water stress treatments was measured as described in the methods. (**C** to **E**) Comparison of transcript levels for TaPP2C6, TaLea, and TaAOS induced by individual ligands in 7-day-old wheat seedlings after a treatment of 50 µM for 12 hours. Statistics were done using a one-way ANOVA and post-hoc Dunnett tests to obtain multiplicity adjusted *P* values for treatment effects relative to mock-treated controls, with *n* = 5 replicates for chlorophyll measurements and *n* = 4 replicates for qRT-PCR studies. Error bars indicate standard error of mean, individual *P* values are shown on the graph. n.s., nonsignificant.

stress on crop yields. We used virtual screening to identify ABA receptor agonists in distinct chemical scaffolds, one of which we modified using scaffold merging to create 3CB and subsequently opabactin, a molecule with high bioactivity and receptor binding affinity. Our cocrystal structure of 3CB:PYL10 reveals a salt bridge to K56 and direct hydrogen bond of its amide to N163, features that likely contribute to the enthalpic binding isotherms observed, and the approximately sevenfold improvement in K_d relative to ABA. Prior sulfonamide ABA mimics do not benefit from the lysine salt bridge, and our work suggests that contacts to this residue are essential for high-affinity binding and pan agonism and should be considered in future ligand design and optimization. Because the sulfonamides in quinabactin, pyrabactin, CB, and AMF4 are positioned proximally to this lysine, it should be possible to generate contacts by lowering the pK_a (where K_a is the acid dissociation constant) of their sulfonamide NH by the installation of adjacent electron withdrawing groups, a strategy previously used to improve sulfonamide dihydrofolate reductase and carbonic anhydrase inhibitors (*54*, *55*). Thus, there are likely many routes to developing more-potent ABA receptor agonists, but our development of a new amide-based scaffold broadens the current chemical space available for manipulating ABA receptor function.

OP's activity in *Arabidopsis* demonstrates that PYL2 contributes to OP's effects on transpiration, in addition to PYR1 and PYL1. In contrast, the effects of sulfonamide agonists require primarily PYR1 and PYL1 (*26*, *33*). We conclude, therefore, that PYL2 is a useful antitranspirant target site in addition to PYR1 and PYL1; this further demonstrates the relevance of subfamily III receptors as targets for the development of synthetic antitranspirants but does not preclude the activation of other receptor types for antitranspirant activity. The development of agonists selective for different receptor subfamilies could help untangle the relative importance of each subfamily to transpiration. Additionally, because increases in both water use and photosynthetic efficiency have been observed by overexpressing subfamily II receptors (*23*, *24*), selective agonists may provide access to the control of distinct water use–related traits not achievable by only activating subfamily III receptors.

The triple mutant *pyr1,pyl1,pyl2* has a relatively modest phenotype. If subfamily III receptors are such critical mediators of OP and sulfonamide effects on transpiration, why is their multilocus loss-of-function phenotype so mild? We do not yet have a complete answer to this question, but one possibility is that other receptors may compensate for the loss of subfamily III receptors, as has been observed by systematic analyses of yeast gene families, which show compensation by paralogs (*56*). Alternatively, it may be that the major effects of OP and sulfonamides require both subfamily II and subfamily III receptors to be present, i.e., different subfamily II receptors than we are currently able to test genetically (*pyl4* and *pyl5*). Given that PYL4 and PYL5 are the most highly expressed subfamily II ABA receptors, this seems unlikely, but future analyses of genetic strains lacking all members of each subfamily should help resolve this. Nonetheless, weighting a protein's relevance as a target for manipulating a biological process on the basis of only genetic studies may be misleading because of compensation. Thus, chemical interrogation of wild-type systems provides a more direct route to assessing the relevance of a protein as a target.

ABA has many effects on plant growth and development in addition to its roles in water relations, seed dormancy, and gene regulation. Although OP mimics ABA's effects, further characterization will be required to establish which of ABA's many responses are activated by OP. It will also be necessary to conduct toxicological assays and regulatory procedures before OP (or any agonist) is suited for agricultural purposes. ABA receptor overexpression can enable increases in water use efficiency without growth penalties in both *Arabidopsis*

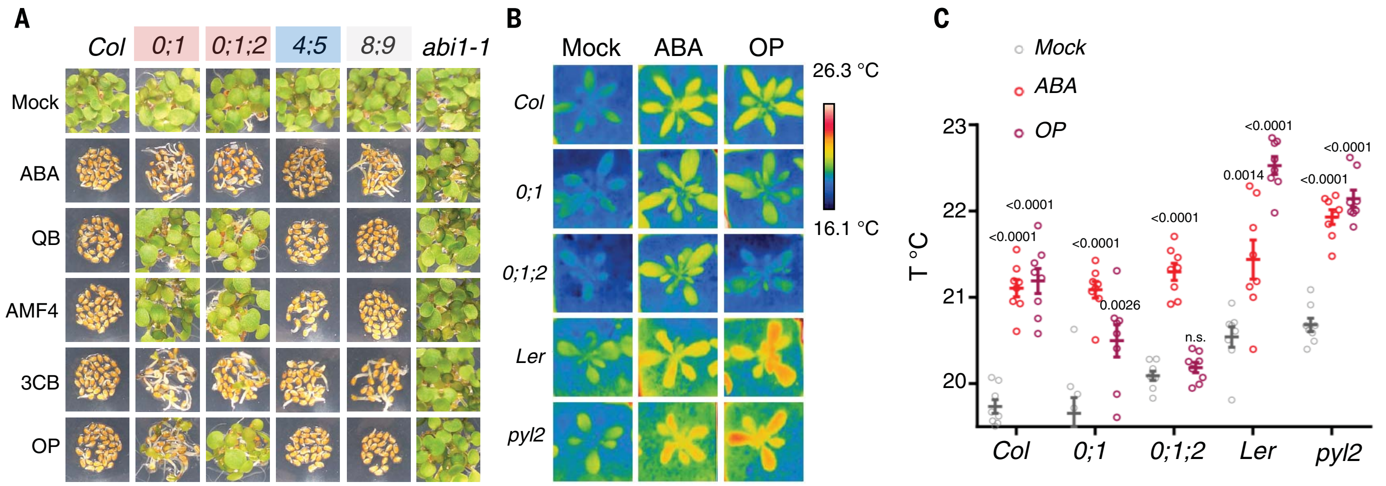


Fig. 7. OP acts primarily through subfamily III receptors and demonstrates that PYL2 contributes to OP's antitranspirant activity. (**A**) Effects of different ABA receptors agonists on seed germination in wild type and different receptor mutant lines. Chemical effects were tested at concentrations approximately two times their IC_{50} values (Fig. 2C) to normalize for differences in agonist potency (ABA, 1200 nM; QB, 1800 nM; AMF4, 1600 nM; 3CB, 440 nM; and OP, 120 nM). Photographs were taken after 4-day incubation under constant illumination. (**B**) Adult Col-0 and different mutant lines were sprayed with 50 μM compounds and imaged by thermography 24 hours after treatment and (**C**) average leaf temperatures for each treatment were quantified. Statistics done using a one-way ANOVA and post-hoc Dunnett tests to obtain multiplicity adjusted *P* values for treatment effects relative to mock-treated controls (n = 8 replicates). Error bars indicate standard error of mean. Individual *P* values for comparisons to mock-treated controls are shown on the graph.

and wheat in controlled-environment studies (*23*, *24*). The application of ABA—which, like OP, inhibits growth at high concentrations—to crops in the field improves yield under modest drought conditions (*57*–*59*), so the growth inhibitory effects of ABA are not an intrinsic limitation to their other benefits. The value of chemical approaches is that they can be deployed on demand, whereas crop genetics cannot be changed dynamically in a growing season. It is likely that both genetics and chemistry will work together for maximal benefit and flexibility (*25*). OP is a new chemical tool for dynamically manipulating plant water use.

Materials and methods summary

Virtual screening of the ZINC collection, a database of commercially available chemical compounds, was carried out using Protein Data Bank (PDB) IDs for PYR1 (3K3K) and PYL9 (3W9R) using Schrodinger's Glide package. Top scoring 0.1% hits were redocked against PYR1 to eliminate ligands not predicted to contact PYR1-K59. 1724 predicted agonists were purchased from Enamine (Ukraine) and tested at 25 μM in a direct receptor activation assay using a pool of recombinant ABA receptors from the three major clades (PYL1, PYL2, PYL4, PYL8, and ZmPYL8). The pooled receptor assay measures agonist dependent receptor mediated inhibition of the protein phosphatase ΔN-HAB1 using a fluorogenic substrate (4-methylumbelliferyl phosphate). This process yielded 25 hits ($IC_{50} \leq$ 25 μM), one of which (3B4) was optimized by chemical synthesis to yield 3CB, a pan agonist. After optimizing one of our hits, we obtained a cocrystal structure of the lead molecule 3CB with the subfamily I ABA receptor PYL10. $PYL10^{25\text{-}183}$:3CB crystals were grown in 33% tacsimate (pH 7.0) and flash frozen using perfluoropolyether oil. Diffraction data was processed in HKL2000, refined in Phenix.refine and validated using MolProbity; the final model was deposited in PDB (ID 6NWC). We next synthesized OP, a derivative of 3CB. 3CB and OP were synthesized from methyl (4-cyanophenyl) acetate over five steps involving (i) palladium catalyzed ortho-halogenation of the nitrile, (ii) suzuki coupling with cyclopropyl boronic acid, (iii) base-catalyzed hydrolysis of the ester, (iv) EDCI/DMAP coupling to methyl 1-aminocyclohexanoate, and (v) subsequent hydrolysis to yield OP (34 to 58% yield). The OP that we synthesized for this paper can be purchased from Kerafast (USA), a distributor of academic research reagents. The AMF4 and tetralone ABA used in our studies were synthesized according to literature protocols, using chiral chromatography to purify the bioactive (+) isomer of tetralone ABA, as previously described. To directly measure agonist-receptor interactions, we conducted isothermal titration calorimetry (ITC) experiments using low-volume Nano ITC (TA Instruments, USA). To determine which residues in PYR1 mediate OP selectivity, we used an established yeast two-hybrid assay to analyze agonist responses in wild-type PYR1, PYR1-V163, and PYR1-V163L mutants. To measure bioactivity of the agonists in adult plants, chemicals were applied to plants as aerosols in a solution supplemented with the wetting agent Silwet-77; leaf temperatures were measured by thermal imaging using a FLIR camera. To profile *Arabidopsis* mutant responses to agonists, we used previously described receptor mutants and a newly constructed *pyr1, pyl1,pyl2* strain that we made by crossing a *pyr1,pyl1,pyl2,pyl4* quadruple mutant strain to wild type (Col-0) and isolating the desired genotype amongst F_3 plants.

REFERENCES AND NOTES

1. S. Jasechko *et al*., Terrestrial water fluxes dominated by transpiration. *Nature* **496**, 347–350 (2013). doi: 10.1038/nature11983; pmid: 23552893
2. M. L. Nuccio, M. Paul, N. J. Bate, J. Cohn, S. R. Cutler, Where are the drought tolerant crops? An assessment of more than two decades of plant biotechnology effort in crop improvement. *Plant Sci.* **273**, 110–119 (2018). doi: 10.1016/j.plantsci.2018.01.020; pmid: 29907303
3. I. Amani, R. A. Fischer, M. P. Reynolds, Canopy temperature depression association with yield of irrigated spring wheat cultivars in a hot climate. *J. Agron. Crop Sci.* **176**, 119–129 (1996). doi: 10.1111/j.1439-037X.1996.tb00454.x
4. R. A. Fischer *et al*., Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Sci.* **38**, 1467–1475 (1998). doi: 10.2135/cropsci1998.0011183X003800060011x
5. J. D. M. Helander, A. S. Vaidya, S. R. Cutler, Chemical manipulation of plant water use. *Bioorg. Med. Chem.* **24**, 493–500 (2016). doi: 10.1016/j.bmc.2015.11.010; pmid: 26612713
6. J. Bruinsma, Ed., *World Agriculture: Towards 2015/2030: An FAO Perspective* (Earthscan, 2003).
7. "Disasters causing billions in agricultural losses, with drought leading the way," Food and Agriculture Organization of the United Nations (15 March 2018); www.fao.org/news/story/en/item/1106977/icode/.
8. J. Frackenpohl *et al*., Use of substituted isoquinolinones, isoquinolindiones, isoquinolintriones and dihydroisoquinolinones or in each case salts thereof as active agents against abiotic stress in plants, U.S. Patent 9173395 (2015); https://patentimages.storage.googleapis.com/85/bf/14/6a7bab7007eff7/US9173395.pdf.
9. S. R. Cutler, S. V. Wendeborn, P. J. Jung, M. D. Lachia, R. Dumeunier, Compounds that induce aba responses, U.S. Patent 20160280651 (2016); https://patentimages.storage.googleapis.com/3c/8e/a5/53ecd16d2042a7/US20160280651A1.pdf.

10. S. R. Cutler, M. Okamoto, Synthetic compounds for vegetative ABA responses, U.S. Patent 9345245 (2016); https://patentimages.storage.googleapis.com/0d/77/e5/b485cd4b65b762/US9345245.pdf.
11. W. Dejonghe, M. Okamoto, S. R. Cutler, Small molecule probes of ABA biosynthesis and signaling. *Plant Cell Physiol.* **59**, 1490–1499 (2018). doi: 10.1093/pcp/pcy126; pmid: 29986078
12. M. D. Lachia *et al*., 2-oxo-3,4-dihydroquinoline compounds as plant growth regulators, U.S. Patent 20180044297 (2018); https://patentimages.storage.googleapis.com/02/4b/ee/fc4a6a0f677484/US20180044297A1.pdf.
13. J. Frackenpohl *et al*., Use of substitute oxo tetrahydroquinoline sulfonamides or salts thereof for raising stress tolerance of plants, U.S. Patent 20170027172 (2017); https://patentimages.storage.googleapis.com/0d/e0/59/3fb3cd32900682/US20170027172A1.pdf.
14. J. Frackenpohl *et al*., Use of substituted dihydrooxindolylsulfonamides, or the salts thereof, for increasing the stress tolerance of plants, U.S. Patent 20160237035 (2016); https://patentimages.storage.googleapis.com/cb/6e/25/842328f2b2cbef/US20160237035A1.pdf.
15. J. Frackenpohl *et al*., Substituted 1-cycloalkyl-2-oxotetrahydroquinolin-6-ylsulfonamides or salts thereof and use thereof to increase stress tolerance in plants, U.S. Patent 20180020662 (2018); https://patentimages.storage.googleapis.com/6d/43/e5/67dd0ad006b0f2/US20180020662A1.pdf.
16. J. Frackenpohl, L. Willms, J. Dittgen, Substituted cyano cycloalkyl penta-2,4-dienes, cyano cycloalkyl pent-2-en-4-ynes, cyano heterocyclyl penta-2,4-dienes and cyano heterocyclyl pent-2-en-4-ynes as active substances, U.S. Patent Application 20170210701A1 (2017); https://patents.google.com/patent/US20170210701A1/en.
17. J. Frackenpohl *et al*., Aryl-and hetarylsulfonamides as active ingredients against abiotic plant stress, U.S. Patent 20110230350 (2011); https://patentimages.storage.googleapis.com/a7/2d/ea/8a60a636e4fdd7/US20110230350A1.pdf.
18. S. R. Cutler, M. D. Lachia, S. V. Wendeborn, C. R. A. Godfrey, D. Sabbadin, Carbamate quinabactin, World Patent 2018017490 (2018); https://patentimages.storage.googleapis.com/65/27/d8/68509d60cc0199/WO2018017490A1.pdf.
19. C. R. A. Godfrey, M. D. Lachia, S. V.Wendeborn, D. Sabbadin, Plant growth regulator compounds, World Patent 2018007217 (2018); https://patentimages.storage.googleapis.com/87/fd/61/816a6c284d84f1/WO2018007217A1.pdf.
20. S. R. Cutler, S. V. Wendeborn, O. Loiseleur, M. D. Lachia, D. Sabbadin, Derivatives of halo quinabactin, U.S. Patent 20180312470 (2018); https://patentimages.storage.googleapis.com/af/59/c9/c06faa2b1d5db6/US20180312470A1.pdf.
21. J. J. Weiner, F. C. Peterson, B. F. Volkman, S. R. Cutler, Structural and functional insights into core ABA signaling. *Curr. Opin. Plant Biol.* **13**, 495–502 (2010). doi: 10.1016/j.pbi.2010.09.007; pmid: 20934900
22. T. Miyakawa, Y. Fujita, K. Yamaguchi-Shinozaki, M. Tanokura, Structure and function of abscisic acid receptors. *Trends Plant Sci.* **18**, 259–266 (2013). doi: 10.1016/j.tplants.2012.11.002; pmid: 23265948
23. Z. Yang *et al*., Leveraging abscisic acid receptors for efficient water use in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6791–6796 (2016). doi: 10.1073/pnas.1601954113; pmid: 27247417
24. R. Mega *et al*., Tuning water-use efficiency and drought tolerance in wheat using abscisic acid receptors. *Nat. Plants* **5**, 153–159 (2019). doi: 10.1038/s41477-019-0361-8; pmid: 30737511
25. M.-J. Cao *et al*., Combining chemical and genetic approaches to increase drought resistance in plants. *Nat. Commun.* **8**, 1183 (2017). doi: 10.1038/s41467-017-01239-3; pmid: 29084945
26. M. Okamoto *et al*., Activation of dimeric ABA receptors elicits guard cell closure, ABA-regulated gene expression, and drought tolerance. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12132–12137 (2013). doi: 10.1073/pnas.1305919110; pmid: 23818638
27. J. Frackenpohl *et al*., Potent analogues of abscisic acid–identifying cyano-cyclopropyl moieties as promising replacements for the cyclohexenone headgroup. *Eur. J. Org. Chem.* **2018**, 1416–1425 (2018). doi: 10.1002/ejoc.201701769
28. J. Frackenpohl *et al*., Insights into the in vitro and in vivo SAR of abscisic acid–exploring unprecedented variations of the side chain via cross-coupling-mediated syntheses. *Eur. J. Org. Chem.* **2018**, 1403–1415 (2018). doi: 10.1002/ejoc.201701687
29. M. Kepka *et al*., Action of natural abscisic acid precursors and catabolites on abscisic acid receptor complexes. *Plant Physiol.* **157**, 2108–2119 (2011). doi: 10.1104/pp.111.182584; pmid: 21976481
30. J. M. Nyangulu *et al*., Synthesis and biological activity of tetralone abscisic acid analogues. *Org. Biomol. Chem.* **4**, 1400–1412 (2006). doi: 10.1039/b509193d; pmid: 16557330
31. A. S. Raghavendra, V. K. Gonugunta, A. Christmann, E. Grill, ABA perception and signalling. *Trends Plant Sci.* **15**, 395–401 (2010). doi: 10.1016/j.tplants.2010.04.006; pmid: 20493758
32. F. Hauser, R. Waadt, J. I. Schroeder, Evolution of abscisic acid synthesis and signaling mechanisms. *Curr. Biol.* **21**, R346–R355 (2011). doi: 10.1016/j.cub.2011.03.015; pmid: 21549957
33. A. S. Vaidya *et al*., A rationally designed agonist defines subfamily IIIA abscisic acid receptors as critical targets for manipulating transpiration. *ACS Chem. Biol.* **12**, 2842–2848 (2017). doi: 10.1021/acschembio.7b00650; pmid: 28949512
34. Y. Yang, A. Costa, N. Leonhardt, R. S. Siegel, J. I. Schroeder, Isolation of a strong Arabidopsis guard cell promoter and its potential as a research tool. *Plant Methods* **4**, 6 (2008). doi: 10.1186/1746-4811-4-6; pmid: 18284694
35. S. Pandey *et al*., Boolean modeling of transcriptome data reveals novel modes of heterotrimeric G-protein action. *Mol. Syst. Biol.* **6**, 372 (2010). doi: 10.1038/msb.2010.28; pmid: 20531402
36. Q. Hao *et al*., Functional mechanism of the abscisic acid agonist pyrabactin. *J. Biol. Chem.* **285**, 28946–28952 (2010). doi: 10.1074/jbc.M110.149005; pmid: 20554531
37. K. Melcher *et al*., Identification and mechanism of ABA receptor antagonism. *Nat. Struct. Mol. Biol.* **17**, 1102–1108 (2010). doi: 10.1038/nsmb.1887; pmid: 20729862
38. F. C. Peterson *et al*., Structural basis for selective activation of ABA receptors. *Nat. Struct. Mol. Biol.* **17**, 1109–1113 (2010). doi: 10.1038/nsmb.1898; pmid: 20729860
39. M. Cao *et al*., An ABA-mimicking ligand that reduces water loss and promotes drought resistance in plants. *Cell Res.* **23**, 1043–1054 (2013). doi: 10.1038/cr.2013.95; pmid: 23835477
40. T. A. Halgren *et al*., Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759 (2004). doi: 10.1021/jm030644s; pmid: 15027866
41. J. J. Irwin, B. K. Shoichet, ZINC—A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182 (2005). doi: 10.1021/ci049714+; pmid: 15667143
42. K. Melcher *et al*., A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. *Nature* **462**, 602–608 (2009). doi: 10.1038/nature08613; pmid: 19898420
43. P. Yin *et al*., Structural insights into the mechanism of abscisic acid signaling by PYL proteins. *Nat. Struct. Mol. Biol.* **16**, 1230–1236 (2009). doi: 10.1038/nsmb.1730; pmid: 19893533
44. K. Miyazono *et al*., Structural basis of abscisic acid signalling. *Nature* **462**, 609–614 (2009). doi: 10.1038/nature08583; pmid: 19855379
45. R. E. Davis, Z. Zhang, B. S. J. Blagg, A scaffold merging approach to Hsp90 C-terminal inhibition: Synthesis and evaluation of a chimeric library. *MedChemComm* **8**, 593–598 (2017). doi: 10.1039/C6MD00377J; pmid: 28533894
46. D. Uraguchi *et al*., A femtomolar-range suicide germination stimulant for the parasitic plant *Striga hermonthica*. *Science* **362**, 1301–1305 (2018). doi: 10.1126/science.aau5445; pmid: 30545887
47. D. Elzinga *et al*., Defining and exploiting hypersensitivity hotspots to facilitate abscisic acid receptor agonist optimization. *ACS Chem. Biol.* **14**, 332–336 (2019). doi: 10.1021/acschembio.8b00955; pmid: 30668093
48. J. Takeuchi *et al*., Designed abscisic acid analogs as antagonists of PYL-PP2C receptor interactions. *Nat. Chem. Biol.* **10**, 477–482 (2014). doi: 10.1038/nchembio.1524; pmid: 24792952
49. E. Barratt *et al*., Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water. *J. Am. Chem. Soc.* **127**, 11827–11834 (2005). doi: 10.1021/ja0527525; pmid: 16104761
50. S. D. Sharrow, M. V. Novotny, M. J. Stone, Thermodynamic analysis of binding between mouse major urinary protein-I and the pheromone 2-sec-butyl-4,5-dihydrothiazole. *Biochemistry* **42**, 6302–6309 (2003). doi: 10.1021/bi026423q; pmid: 12755635
51. R. J. Bingham *et al*., Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein. *J. Am. Chem. Soc.* **126**, 1675–1681 (2004). doi: 10.1021/ja038461i; pmid: 14871097
52. E. Freire, Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* **13**, 869–874 (2008). doi: 10.1016/j.drudis.2008.07.005; pmid: 18703160
53. J. E. Ladbury, G. Klebe, E. Freire, Adding calorimetric data to decision making in lead discovery: A hot tip. *Nat. Rev. Drug Discov.* **9**, 23–27 (2010). doi: 10.1038/nrd3054; pmid: 19960014
54. V. Dudutienė *et al*., Discovery and characterization of novel selective inhibitors of carbonic anhydrase IX. *J. Med. Chem.* **57**, 9435–9446 (2014). doi: 10.1021/jm501003k; pmid: 25358084
55. T. L. Lemke, D. A. Williams, *Foye's Principles of Medicinal Chemistry* (Lippincott Williams & Wilkins, 2008).
56. G. Diss *et al*., Gene duplication can impart fragility, not robustness, in the yeast protein interaction network. *Science* **355**, 630–634 (2017). doi: 10.1126/science.aai7685; pmid: 28183979
57. C. Travaglia *et al*., Exogenous ABA increases yield in field-grown wheat with moderate water restriction. *J. Plant Growth Regul.* **29**, 366–374 (2010). doi: 10.1007/s00344-010-9147-y
58. X. Zhang *et al*., Improving winter wheat performance by foliar spray of ABA and FA under water deficit conditions. *J. Plant Growth Regul.* **35**, 83–96 (2016). doi: 10.1007/s00344-015-9509-6
59. T. C. de Souza *et al*., ABA application to maize hybrids contrasting for drought tolerance: Changes in water parameters and in antioxidant enzyme activity. *Plant Growth Regul.* **73**, 205–217 (2014). doi: 10.1007/s10725-013-9881-9

ACKNOWLEDGMENTS

Funding: NSF IOS (grant 1656890) and Syngenta Crop Protection AG to S.R.C. JST PRESTO (JPMJPR15Q5), KAKENHI (17H05009), and the Joint Research Program of Arid Land Research Center, Tottori University (30C2007) to M.O. **Author contributions:** A.S.V. designed and synthesized 3CB and OP, with technical assistance from B.K. and S.B., and performed *Arabidopsis* PP2C assays, seed germination assays, luciferase imaging, and thermal imaging experiments. J.D.M.H. performed virtual screening and hit characterization. F.C.P. performed crystallographic studies in coordination with B.F.V. D.E. performed isothermal titration calorimetry studies. R.M. performed plasmid vector construction for PP2C assay. W.D. performed SEM studies, SAI, root growth, and greening measurements. A.K. and J.D.M.H. conducted small molecule screens of virtual screening hits. S.-Y.P. performed yeast two-hybrid assays. Z.X. and S.-Y.P. purified and characterized *Arabidopsis* PYR/PYL proteins. M.O. identified, cloned, and expressed wheat ABA receptors and performed wheat PP2C assays, qRT-PCR, and drought stress assays. J.T. and Y.T. synthesized tetralone ABA for comparison in various assays. S.R.C. conceived of the research, and A.S.V. and S.R.C. wrote the paper with contributions from all authors. **Competing interests:** S.R.C., J.D.M.H., and A.S.V. are inventors on a UC-owned patent application (62/691,534) covering OP and related structures. **Data and materials availability:** All data are presented in the main text and supplementary materials, the OP synthesized for this study can be obtained from Kerafast, and other biological materials may be obtained by contacting S.R.C. (*Arabidopsis*) or M.O. (wheat). The atomic coordinates and structure factor files for the PYL103CB structure have been deposited in Protein Data Bank (ID 6NWC).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/eaaw8848/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S6
NMR Spectra
References (*60*–*75*)

View/request a protocol for this paper from *Bio-protocol*.

6 February 2019; accepted 11 September 2019
10.1126/science.aaw8848

RESEARCH ARTICLE

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2]*, Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5]*†

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

There is growing concern that algorithms may reproduce racial and gender disparities via the people building them or through the data used to train them (*1*–*3*). Empirical work is increasingly lending support to these concerns. For example, job search ads for highly paid positions are less likely to be presented to women (*4*), searches for distinctively Black-sounding names are more likely to trigger ads for arrest records (*5*), and image searches for professions such as CEO produce fewer images of women (*6*). Facial recognition systems increasingly used in law enforcement perform worse on recognizing faces of women and Black individuals (*7*, *8*), and natural language processing algorithms encode language in gendered ways (*9*).

Empirical investigations of algorithmic bias, though, have been hindered by a key constraint: Algorithms deployed on large scales are typically proprietary, making it difficult for independent researchers to dissect them. Instead, researchers must work "from the outside," often with great ingenuity, and resort to clever workarounds such as audit studies. Such efforts can document disparities, but understanding how and why they arise—much less figuring out what to do about them—is difficult without greater access to the algorithms themselves. Our understanding of a mechanism therefore typically relies on theory or exercises with researcher-created algorithms (*10*–*13*). Without an algorithm's training data, objective function, and prediction methodology, we can only guess as to the actual mechanisms for the important algorithmic disparities that arise.

In this study, we exploit a rich dataset that provides insight into a live, scaled algorithm deployed nationwide today. It is one of the largest and most typical examples of a class of commercial risk-prediction tools that, by industry estimates, are applied to roughly 200 million people in the United States each year. Large health systems and payers rely on this algorithm to target patients for "high-risk care management" programs. These programs seek to improve the care of patients with complex health needs by providing additional resources, including greater attention from trained providers, to help ensure that care is well coordinated. Most health systems use these programs as the cornerstone of population health management efforts, and they are widely considered effective at improving outcomes and satisfaction while reducing costs (*14*–*17*). Because the programs are themselves expensive—with costs going toward teams of dedicated nurses, extra primary care appointment slots, and other scarce resources—health systems rely extensively on algorithms to identify patients who will benefit the most (*18*, *19*).

Identifying patients who will derive the greatest benefit from these programs is a challenging causal inference problem that requires estimation of individual treatment effects. To solve this problem, health systems make a key assumption: Those with the greatest care needs will benefit the most from the program. Under this assumption, the targeting problem becomes a pure prediction policy problem (*20*). Developers then build algorithms that rely on past data to build a predictor of future health care needs.

Our dataset describes one such typical algorithm. It contains both the algorithm's predictions as well as the data needed to understand its inner workings: that is, the underlying ingredients used to form the algorithm (data, objective function, etc.) and links to a rich set of outcome data. Because we have the inputs, outputs, and eventual outcomes, our data allow us a rare opportunity to quantify racial disparities in algorithms and isolate the mechanisms by which they arise. It should be emphasized that this algorithm is not unique. Rather, it is emblematic of a generalized approach to risk prediction in the health sector, widely adopted by a range of for- and non-profit medical centers and governmental agencies (*21*).

Our analysis has implications beyond what we learn about this particular algorithm. First, the specific problem solved by this algorithm has analogies in many other sectors: The predicted risk of some future outcome (in our case, health care needs) is widely used to target policy interventions under the assumption that the treatment effect is monotonic in that risk, and the methods used to build the algorithm are standard. Mechanisms of bias uncovered in this study likely operate elsewhere. Second, even beyond our particular finding, we hope that this exercise illustrates the importance, and the large opportunity, of studying algorithmic bias in health care, not just as a model system but also in its own right. By any standard—e.g., number of lives affected, life-and-death consequences of the decision—health is one of the most important and widespread social sectors in which algorithms are already used at scale today, unbeknownst to many.

Data and analytic strategy

Working with a large academic hospital, we identified all primary care patients enrolled in risk-based contracts from 2013 to 2015. Our primary interest was in studying differences between White and Black patients. We formed race categories by using hospital records, which are based on patient self-reporting. Any patient who identified as Black was considered to be Black for the purpose of this analysis. Of the remaining patients, those who self-identified as races other than White (e.g., Hispanic) were so considered (data on these patients are presented in table S1 and fig. S1 in the supplementary materials). We considered all remaining patients to be White. This approach allowed us to study one particular racial difference of social and historical interest between patients who self-identified as Black and patients who self-identified as White without another race or ethnicity; it has the disadvantage of not allowing for the study of intersectional racial

[1]School of Public Health, University of California, Berkeley, Berkeley, CA, USA. [2]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, USA. [3]Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA. [4]Mongan Institute Health Policy Center, Massachusetts General Hospital, Boston, MA, USA. [5]Booth School of Business, University of Chicago, Chicago, IL, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: sendhil.mullainathan@chicagobooth.edu

and ethnic identities. Our main sample thus consisted of (i) 6079 patients who self-identified as Black and (ii) 43,539 patients who self-identified as White without another race or ethnicity, whom we observed over 11,929 and 88,080 patient-years, respectively (1 patient-year represents data collected for an individual patient in a calendar year). The sample was 71.2% enrolled in commercial insurance and 28.8% in Medicare; on average, 50.9 years old; and 63% female (Table 1).

For these patients, we obtained algorithmic risk scores generated for each patient-year. In the health system we studied, risk scores are generated for each patient during the enrollment period for the system's care management program. Patients above the 97th percentile are automatically identified for enrollment in the program. Those above the 55th percentile are referred to their primary care physician, who is provided with contextual data about the patients and asked to consider whether they would benefit from program enrollment.

Many existing metrics of algorithmic bias may apply to this scenario. Some definitions focus on calibration [i.e., whether the realized value of some variable of interest Y matches the risk score R (*2*, *22*, *23*)]; others on statistical parity of some decision D influenced by the algorithm (*10*); and still others on balance of average predictions, conditional on the realized outcome (*22*). Given this multiplicity and the growing recognition that not all conditions can be simultaneously satisfied (*3*, *10*, *22*), we focus on metrics most relevant to the real-world use of the algorithm, which are related to calibration bias [formally, comparing Blacks B and Whites W, $E[Y|R,W] = E[Y|R,B]$ indicates the absence of bias (here, E is the expectation operator)]. The algorithm's stated goal is to predict complex health needs for the purpose of targeting an intervention that manages those needs. Thus, we compare the algorithmic risk score for patient i in year t ($R_{i,t}$), formed on the basis of claims data $X_{i,(t-1)}$ from the prior year, to data on patients' realized health $H_{i,t}$, assessing how well the algorithmic risk score is calibrated across race for health outcomes $H_{i,t}$. We also ask how well the algorithm is calibrated for costs $C_{i,t}$.

To measure H, we link predictions to a wide range of outcomes in electronic health record data, including all diagnoses (in the form of International Classification of Diseases codes) as well as key quantitative laboratory studies and vital signs capturing the severity of chronic illnesses. To measure C, we link predictions to insurance claims data on utilization, including outpatient and emergency visits, hospitalizations, and health care costs. These data, and the rationale for the specific measures of H used in this study, are described in more detail in the supplementary materials.

Health disparities conditional on risk score

We begin by calculating an overall measure of health status, the number of active chronic conditions [or "comorbidity score," a metric used extensively in medical research (*24*) to provide a comprehensive view of a patient's health (*25*)] by race, conditional on algorithmic risk score. Fig. 1A shows that, at the same level of algorithm-predicted risk, Blacks have significantly more illness burden than Whites. We can quantify these differences by choosing one point on the x axis that corresponds to a very-high-risk group (e.g., patients at the 97th percentile of risk score, at which patients are auto-identified for program enrollment), where Blacks have 26.3% more chronic illnesses than Whites (4.8 versus 3.8 distinct conditions; $P < 0.001$).

What do these prediction differences mean for patients? Algorithm scores are a key input to decisions about future enrollment in a care coordination program. So as we might expect, with less-healthy Blacks scored at similar risk scores to more-healthy Whites, we find evidence

Table 1. Descriptive statistics on our sample, by race. BP, blood pressure; LDL, low-density lipoprotein.

	White	Black
n (patient-years)	88,080	11,929
n (patients)	43,539	6079
Demographics		
Age	51.3	48.6
Female (%)	62	69
Care management program		
Algorithm score (percentile)	50	52
Race composition of program (%)	81.8	18.2
Care utilization		
Actual cost	$7540	$8442
Hospitalizations	0.09	0.13
Hospital days	0.50	0.78
Emergency visits	0.19	0.35
Outpatient visits	4.94	4.31
Mean biomarker values		
HbA1c (%)	5.9	6.4
Systolic BP (mmHg)	126.6	130.3
Diastolic BP (mmHg)	75.5	75.7
Creatinine (mg/dl)	0.89	0.98
Hematocrit (%)	40.7	37.8
LDL (mg/dl)	103.4	103.0
Active chronic illnesses (comorbidities)		
Total number of active illnesses	1.20	1.90
Hypertension	0.29	0.44
Diabetes, uncomplicated	0.08	0.22
Arrythmia	0.09	0.08
Hypothyroid	0.09	0.05
Obesity	0.07	0.18
Pulmonary disease	0.07	0.11
Cancer	0.07	0.06
Depression	0.06	0.08
Anemia	0.05	0.10
Arthritis	0.04	0.04
Renal failure	0.03	0.07
Electrolyte disorder	0.03	0.05
Heart failure	0.03	0.05
Psychosis	0.03	0.05
Valvular disease	0.03	0.02
Stroke	0.02	0.03
Peripheral vascular disease	0.02	0.02
Diabetes, complicated	0.02	0.07
Heart attack	0.01	0.02
Liver disease	0.01	0.02

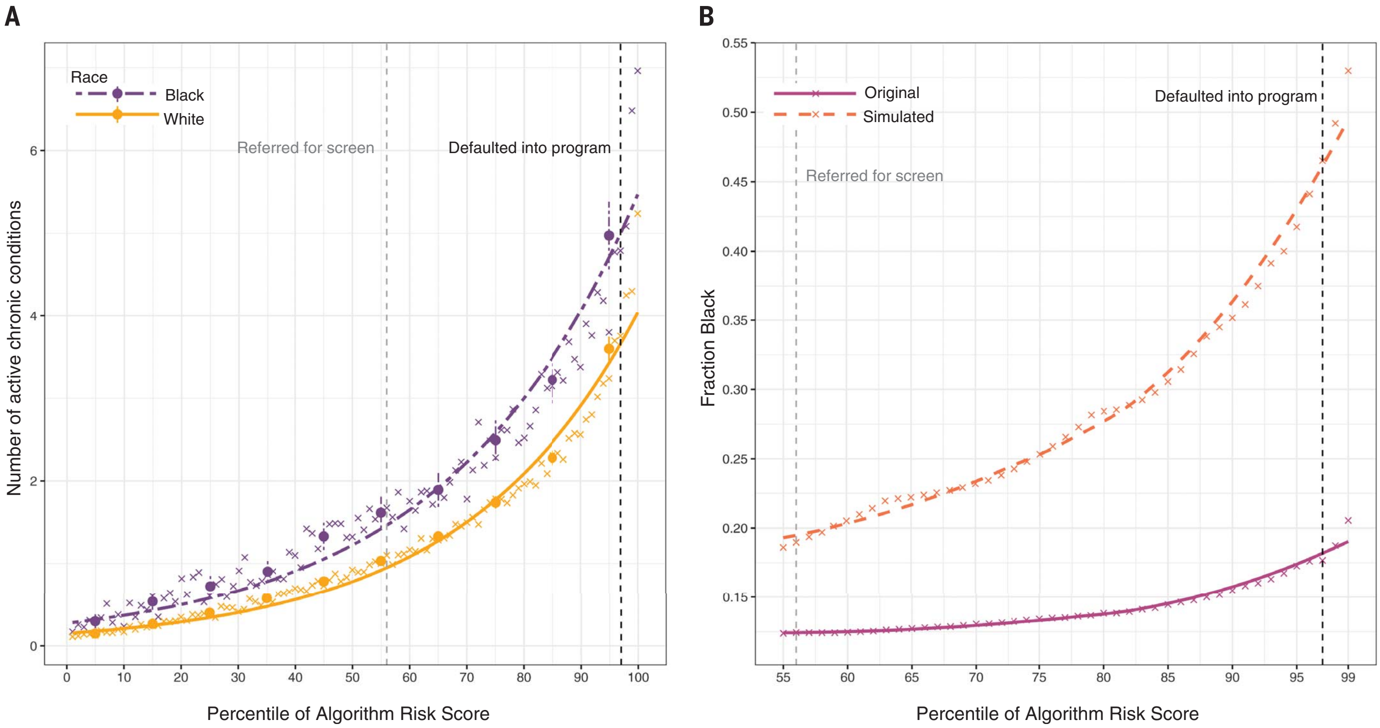


Fig. 1. Number of chronic illnesses versus algorithm-predicted risk, by race. (**A**) Mean number of chronic conditions by race, plotted against algorithm risk score. (**B**) Fraction of Black patients at or above a given risk score for the original algorithm ("original") and for a simulated scenario that removes algorithmic bias ("simulated": at each threshold of risk, defined at a given percentile on the *x* axis, healthier Whites above the threshold are replaced with less healthy Blacks below the threshold, until the marginal patient is equally healthy). The × symbols show risk percentiles by race; circles show risk deciles with 95% confidence intervals clustered by patient. The dashed vertical lines show the auto-identification threshold (the black line, which denotes the 97th percentile) and the screening threshold (the gray line, which denotes the 55th percentile).

of substantial disparities in program screening. We quantify this by simulating a counterfactual world with no gap in health conditional on risk. Specifically, at some risk threshold α, we identify the supramarginal White patient (i) with $R_i > \alpha$ and compare this patient's health to that of the inframarginal Black patient (j) with $R_j < \alpha$. If $H_i > H_j$, as measured by number of chronic medical conditions, we replace the (healthier, but supramarginal) White patient with the (sicker, but inframarginal) Black patient. We repeat this procedure until $H_i = H_j$, to simulate an algorithm with no predictive gap between Blacks and Whites. Fig. 1B shows the results: At all risk thresholds α above the 50th percentile, this procedure would increase the fraction of Black patients. For example, at α = 97th percentile, among those auto-identified for the program, the fraction of Black patients would rise from 17.7 to 46.5%.

We then turn to a more multidimensional picture of the complexity and severity of patients' health status, as measured by biomarkers that index the severity of the most common chronic illnesses in our sample (as shown in Table 1). This allows us to identify patients who might derive a great deal of benefit from care management programs—e.g., patients with severe diabetes who are at risk of catastrophic complications if they do not lower their blood sugar (*18*, *26*). (The materials and methods section describes several experiments to rule out a large effect of the program on these health measures in year t; had there been such an effect, we could not easily use the measures to assess the accuracy of the algorithm's predictions on health, because the program is allocated as a function of algorithm score.) Across all of these important markers of health needs—severity of diabetes, high blood pressure, renal failure, cholesterol, and anemia—we find that Blacks are substantially less healthy than Whites at any level of algorithm predictions, as shown in Fig. 2. Blacks have more-severe hypertension, diabetes, renal failure, and anemia, and higher cholesterol. The magnitudes of these differences are large: For example, differences in severity of hypertension (systolic pressure: 5.7 mmHg) and diabetes [glycated hemoglobin (HbA1c): 0.6%] imply differences in all-cause mortality of 7.6% (*27*) and 30% (*28*), respectively, calculated using data from clinical trials and longitudinal studies.

Mechanism of bias

An unusual aspect of our dataset is that we observe the algorithm's inputs and outputs as well as its objective function, providing us a unique window into the mechanisms by which bias arises. In our setting, the algorithm takes in a large set of raw insurance claims data $X_{i,t-1}$ (features) over the year t – 1: demographics (e.g., age, sex), insurance type, diagnosis and procedure codes, medications, and detailed costs. Notably, the algorithm specifically excludes race.

The algorithm uses these data to predict $Y_{i,t}$ (i.e., the label). In this instance, the algorithm takes total medical expenditures (for simplicity, we denote "costs" C_t) in year t as the label. Thus, the algorithm's prediction on health needs is, in fact, a prediction on health costs.

As a first check on this potential mechanism of bias, we calculate the distribution of realized costs C versus predicted costs R. By this metric, one could call the algorithm unbiased. Fig. 3A shows that, at every level of algorithm-predicted risk, Blacks and Whites have (roughly) the same costs the following year. In other words, the algorithm's predictions are well calibrated across races. For example, at the median risk score, Black patients had costs of \$5147 versus \$4995 for Whites (U.S. dollars); in the top 5% of algorithm-predicted risk, costs were \$35,541 for Blacks versus \$34,059 for Whites.

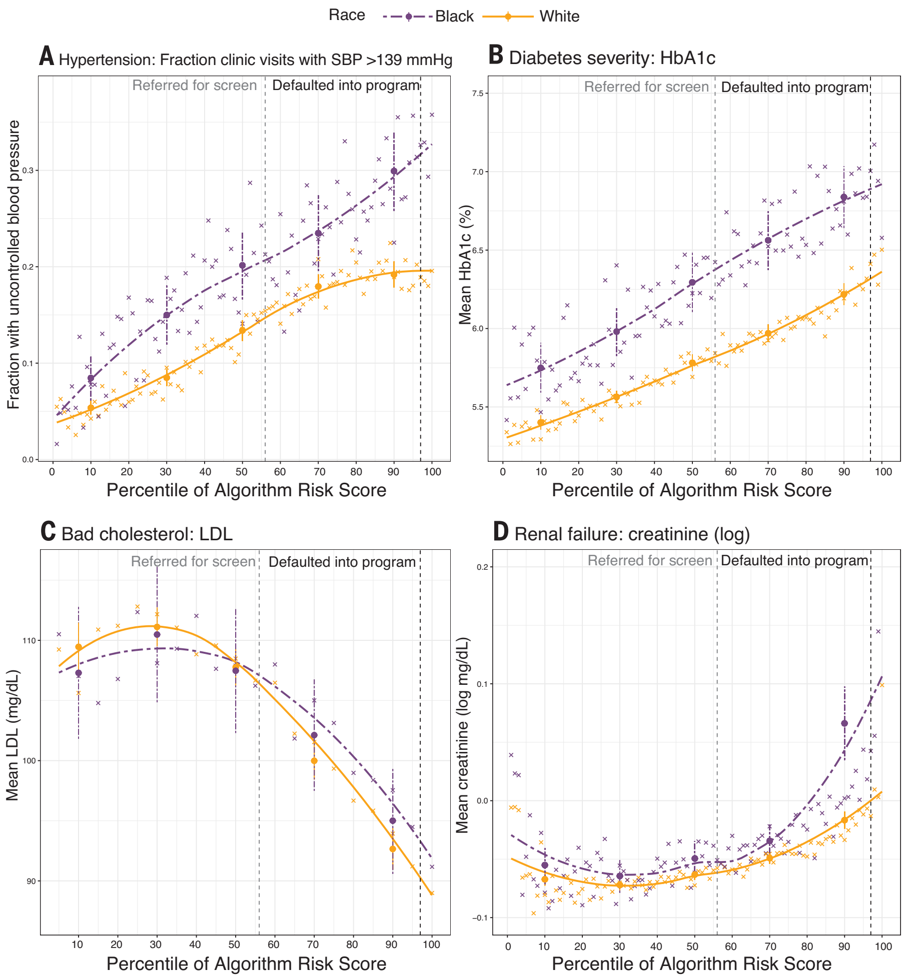
Because these programs are used to target patients with high costs, these results are largely inconsistent with algorithmic bias, as measured by calibration: Conditional on risk score, predictions do not favor Whites or Blacks anywhere in the risk distribution.

To summarize, we find substantial disparities in health conditional on risk but little disparity in costs. On the one hand, this is surprising: Health care costs and health needs are highly correlated, as sicker patients need and receive more care, on average. On the other hand, there are many opportunities for a wedge to creep in between needing health care and receiving health care—and crucially, we find that wedge to be correlated with race, as shown in Fig. 3B. At a given level of health (again measured by number of chronic illnesses), Blacks generate lower costs than Whites—on average, $1801 less per year, holding constant the number of chronic illnesses (or $1144 less, if we instead hold constant the specific individual illnesses that contribute to the sum). Table S2 also shows that Black patients generate very different kinds of costs: for example, fewer inpatient surgical and outpatient specialist costs, and more costs related to emergency visits and dialysis. These results suggest that the driving force behind the bias we detect is that Black patients generate lesser medical expenses, conditional on health, even when we account for specific comorbidities. As a result, accurate prediction of costs necessarily means being racially biased on health.

How might these disparities in cost arise? The literature broadly suggests two main potential channels. First, poor patients face substantial barriers to accessing health care, even when enrolled in insurance plans. Although the population we study is entirely insured, there are many other mechanisms by which poverty can lead to disparities in use of health care: geography and differential access to transportation, competing demands from jobs or child care, or knowledge of reasons to seek care (*29–31*). To the extent that race and socioeconomic status are correlated, these factors will differentially affect Black patients. Second, race could affect costs directly via several channels: direct ("taste-based") discrimination, changes to the doctor–patient relationship, or others. A recent trial randomly assigned Black patients to a Black or White primary care provider and found significantly higher uptake of recommended preventive care when the provider was Black (*32*). This is perhaps the most rigorous demonstration of this effect, and it fits with a larger literature on potential mechanisms by which race can affect health care directly. For example, it has long been documented that Black patients have reduced trust in the health care system (*33*), a fact that some studies trace to the revelations of the Tuskegee study and other adverse experiences (*34*). A substantial literature in psychology has documented physicians' differential perceptions of Black patients, in terms of intelligence, affiliation (*35*), or pain tolerance (*36*). Thus, whether it is communication, trust, or bias, something about the interactions of Black patients with the health care system itself leads to reduced use of health care. The collective effect of these many channels is to lower health spending substantially for Black patients, conditional on need—a finding that has been appreciated for at least two decades (*37*).

Problem formulation

Our findings highlight the importance of the choice of the label on which the algorithm is trained. On the one hand, the algorithm manufacturer's choice to predict future costs is reasonable: The program's goal, at least in part, is



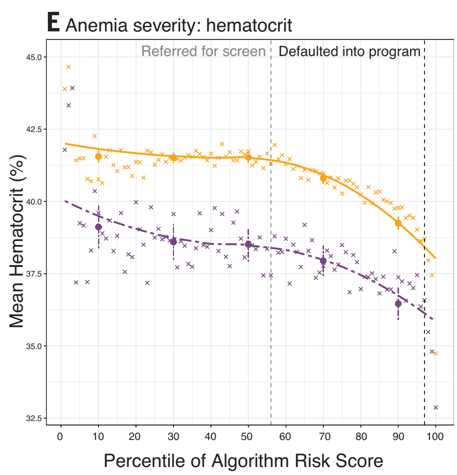


Fig. 2. Biomarkers of health versus algorithm-predicted risk, by race. (**A** to **E**) Racial differences in a range of biological measures of disease severity, conditional on algorithm risk score, for the most common diseases in the population studied. The × symbols show risk percentiles by race, except in (C) where they show risk ventiles; circles show risk quintiles with 95% confidence intervals clustered by patient. The *y* axis in (D) has been trimmed for readability, so the highest percentiles of values for Black patients are not shown. The dashed vertical lines show the auto-identification threshold (black line: 97th percentile) and the screening threshold (gray line: 55th percentile).

to reduce costs, and it stands to reason that patients with the greatest future costs could have the greatest benefit from the program. As noted in the supplementary materials, the manufacturer is not alone. Although the details of individual algorithms vary, the cost label reflects the industry-wide approach. For example, the Society of Actuaries's comprehensive evaluation of the 10 most widely used algorithms, including the particular algorithm we study, used cost prediction as its accuracy metric (*21*). As noted in the report, the enthusiasm for cost prediction is not restricted to industry: Similar algorithms are developed and used by non-profit hospitals, academic groups, and governmental agencies, and are often described in academic literature on targeting population health interventions (*18*, *19*).

On the other hand, future cost is by no means the only reasonable choice. For example, the evidence on care management programs shows that they do not operate to reduce costs globally. Rather, these programs primarily work to prevent acute health decompensations that lead to catastrophic health care utilization (indeed, they actually work to increase other categories of costs, such as primary care and home health assistance; see table S2). Thus avoidable future costs, i.e., those related to emergency visits and hospitalizations, could be a useful label to predict. Alternatively, rather than predicting costs at all, we could simply predict a measure of health; e.g., the number of active chronic health conditions. Because the program ultimately operates to improve the management of these conditions, patients with the most encounters related to them could also be a promising group on which to deploy preventative interventions.

The dilemma of which label to choose relates to a growing literature on "problem formulation" in data science: the task of turning an often amorphous concept we wish to predict into a concrete variable that can be predicted in a given dataset (*38*). Problems in health seem particularly challenging: Health is, by nature, holistic and multidimensional, and there is no single, precise way to measure it. Health care costs, though well measured and readily available in insurance claims data, are also the result of a complex aggregation process with a number of distortions due to structural inequality, incentives, and inefficiency. So although the choice of label is perhaps the single most important decision made in the development of a prediction algorithm, in our setting and in many others, there is often a confusingly large array of different options, each with its own profile of costs and benefits.

Experiments on label choice

Through a series of experiments with our dataset, we can gain some insight into how label choice affects both predictive performance and racial bias. We develop three new predictive algorithms, all trained in the same way, to predict the following outcomes: total cost in year *t* (this tailors cost predictions to our own dataset rather than the national training set), avoidable cost in year *t* (due to emergency visits and hospitalizations), and health in year *t* (measured by the number of chronic conditions that flare up in that year). We train all models in a random ⅔ training set and show all results only from the ⅓ holdout set. Furthermore, as with the original algorithm, we exclude race from the feature set (more details are in the materials and methods).

Table 2 shows the results of these experiments. The first finding is that all algorithms perform reasonably well for predicting not only the outcome on which they were trained but also the other outcomes: The concentration of realized outcomes in those at or above the 97th percentile is notably similar for all algorithms across all outcomes. The largest difference in performance across algorithms is seen for cost prediction: Of all costs in the holdout set, the fraction generated by those at or above the 97th percentile is 16.5% for the cost predictor versus 12.1% for the predictor

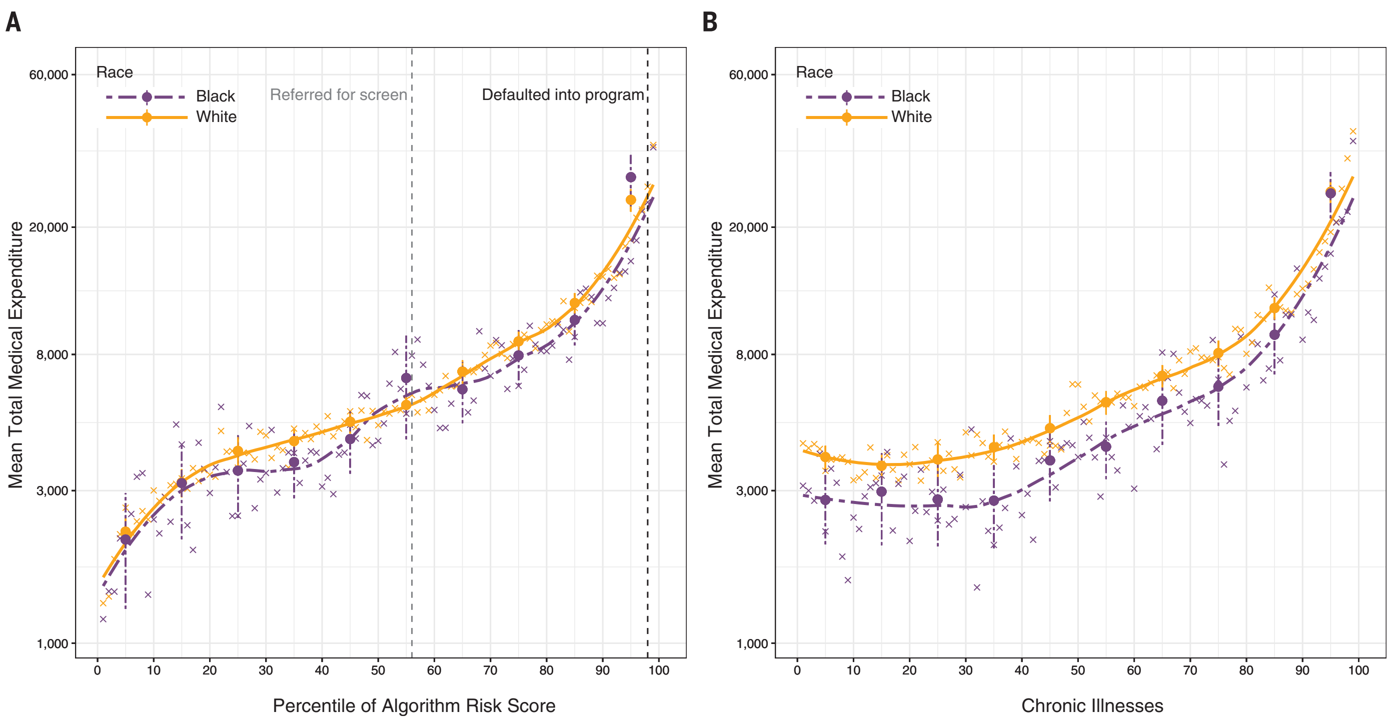


Fig. 3. Costs versus algorithm-predicted risk, and costs versus health, by race. (**A**) Total medical expenditures by race, conditional on algorithm risk score. The dashed vertical lines show the auto-identification threshold (black line: 97th percentile) and the screening threshold (gray line: 55th percentile). (**B**) Total medical expenditures by race, conditional on number of chronic conditions. The × symbols show risk percentiles; circles show risk deciles with 95% confidence intervals clustered by patient. The *y* axis uses a log scale.

of chronic conditions. We then test for label choice bias, defined analogously to calibration bias above: For two algorithms trained to predict Y and Y', and using a threshold τ indexing a (similarly sized) high-risk group, we would test $p[B|R > \tau] = p[B|R' > \tau]$ (here, p denotes probability and B represents Black patients).

We find that the racial composition of this highest-risk group varies far more across algorithms: The fraction of Black patients at or above these risk levels ranges from 14.1% for the cost predictor to 26.7% for the predictor of chronic conditions. Thus, although there could be many reasonable choices of label—all predictions are highly correlated, and any could be justified as a measure of patients' likely benefit from the program—they have markedly different implications in terms of bias, with nearly twofold variation in composition of Black patients in the highest-risk groups.

Relation to human judgment

As noted above, the algorithm is not used for program enrollment decisions in isolation. Rather, it is used as a screening tool, in part to alert primary care doctors to high-risk patients. Specifically, for patients at or above a certain level of predicted risk (the 55th percentile), doctors are presented with contextual information from patients' electronic health records and insurance claims and are prompted to consider enrolling them in the program. Thus, realized enrollment decisions largely reflect how doctors respond to algorithmic predictions, along with other administrative factors related to eligibility (for instance, primary care practice site, residence outside of a nursing home, and continual enrollment in an insurance plan).

Table 3 shows statistics on those enrolled in the program, accounting for 1.3% of observations in our sample: The enrolled individuals are 19.2% Black (versus 11.9% Black in our entire sample) and account for 2.9% of all costs and 3.3% of all active chronic conditions in the population as a whole. We then perform four counterfactual simulations to put these numbers in context; naturally, these simulations use only observable factors, not the many unobserved administrative and human factors that also affect enrollment. First, we calculate the realized program enrollment rate within each percentile of the original algorithm's predicted risk bins and randomly sample patients in each bin for enrollment. This simulation, which mimics "race-blind" enrollment conditional on algorithm score, would yield an enrolled population that is 18.3% Black (versus 19.2% observed; $P = 0.8348$). Second, rather than randomly sampling, we sample those with the highest predicted number of active chronic conditions within a risk bin (using our experimental algorithm described above); this would yield a population that is 26.9% Black. Finally, we compare this to simply assigning those with the highest predicted costs, or the highest number of active chronic conditions, to the program (also using our own algorithms detailed above), which would yield 17.2 and 29.2% Black patients, respectively. Thus, although doctors do redress a small part of the algorithm's bias, they do so far less than an algorithm trained on a different label.

Discussion

Bias attributable to label choice—the difference between some unobserved optimal prediction and the prediction of an algorithm trained on an observed label—is a useful framework through which to understand bias in algorithms, both

Table 2. Performance of predictors trained on alternative labels. For each new algorithm, we show the label on which it was trained (rows) and the concentration of a given outcome of interest (columns) at or above the 97th percentile of predicted risk. We also show the fraction of Black patients in each group.

Algorithm training label	Concentration in highest-risk patients (SE)						Fraction of Black patients in group with highest risk (SE)	
	Total costs		Avoidable costs		Active chronic conditions			
Total costs	0.165	(0.003)	0.187	(0.003)	0.105	(0.002)	0.141	(0.003)
Avoidable costs	0.142	(0.003)	0.215	(0.003)	0.130	(0.003)	0.210	(0.003)
Active chronic conditions	0.121	(0.003)	0.182	(0.003)	0.148	(0.003)	0.267	(0.003)
Best-to-worst difference	0.044		0.033		0.043		0.126	

Table 3. Doctors' decisions versus algorithmic predictions. For those enrolled in the high-risk care management program (1.3% of our sample), we first show the fraction of the population that is Black, as well as the fraction of all costs and chronic conditions accounted for by these observations. We also show these quantities for four alternative program enrollment rules, which we simulate in our dataset (using the holdout set when we use our experimental predictors). We first calculate the program enrollment rate within each percentile bin of predicted risk from the original algorithm and either (i) randomly sample patients or (ii) sample those with the highest predicted number of active chronic conditions within a bin and assign them to the program. The resultant values are then compared with values obtained by simply assigning the aforementioned 1.3% of our sample with (iii) the highest predicted cost or (iv) the highest number of active chronic conditions to the program.

Population	Fraction Black (SE)		Fraction of all costs (SE)		Fraction of all active chronic conditions (SE)	
Observed program enrollment (1.3%)	0.192	(0.003)	0.029	(0.001)	0.033	(0.001)
Simulated alternative enrollment rules						
Random, in predicted-cost bin	0.183	(0.003)	0.044	(0.002)	0.034	(0.001)
Predicted health, in predicted-cost bin	0.269	(0.003)	0.044	(0.002)	0.064	(0.002)
Highest predicted cost	0.172	(0.003)	0.100	(0.002)	0.047	(0.002)
Worst predicted health	0.292	(0.004)	0.067	(0.002)	0.076	(0.002)

in the health sector and further afield. This is because labels are often measured with errors that reflect structural inequalities (*39*). Within the health sector, using mortality or readmission rates to measure hospital performance penalizes those serving poor or non-White populations (*40*, *41*). Outside of the health arena, credit-scoring algorithms predict outcomes related to income, thus incorporating disparities in employment and salary (*2*). Policing algorithms predict measured crime, which also reflects increased scrutiny of some groups (*42*). Hiring algorithms predict employment decisions or supervisory ratings, which are affected by race and gender biases (*43*). Even retail algorithms, which set pricing for goods at the national level, penalize poorer households, which are subjected to increased prices as a result (*44*).

This mechanism of bias is particularly pernicious because it can arise from reasonable choices: Using traditional metrics of overall prediction quality, cost seemed to be an effective proxy for health yet still produced large biases. After completing the analyses described above, we contacted the algorithm manufacturer for an initial discussion of our results. In response, the manufacturer independently replicated our analyses on its national dataset of 3,695,943 commercially insured patients. This effort confirmed our results—by one measure of predictive bias calculated in their dataset, Black patients had 48,772 more active chronic conditions than White patients, conditional on risk score—illustrating how biases can indeed arise inadvertently.

To resolve the issue, we began to experiment with solutions together. As a first step, we suggested using the existing model infrastructure—sample, predictors (excluding race, as before), training process, and so forth—but changing the label: Rather than future cost, we created an index variable that combined health prediction with cost prediction. This approach reduced the number of excess active chronic conditions in Blacks, conditional on risk score, to 7758, an 84% reduction in bias. Building on these results, we are establishing an ongoing (unpaid) collaboration to convert the results of Table 3 into a better, scaled predictor of multidimensional health measures, with the goal of rolling these improvements out in a future round of algorithm development. Of course, our experience may not be typical of all algorithm developers in this sector. But because the manufacturer of the algorithm we study is widely viewed as an industry leader in data and analytics, we are hopeful that this endeavor will prompt other manufacturers to implement similar fixes.

These results suggest that label biases are fixable. Changing the procedures by which we fit algorithms (for instance, by using a new statistical technique for decorrelating predictors with race or other similar solutions) is not required. Rather, we must change the data we feed the algorithm—specifically, the labels we give it. Producing new labels requires deep understanding of the domain, the ability to identify and extract relevant data elements, and the capacity to iterate and experiment. But there is precedent for all of these functions in the literature and, more concretely, in the private companies that invest heavily in developing new and improved labels to predict factors such as consumer behavior (*45*). In addition, although health—as well as criminal justice, employment, and other socially important areas—presents substantial challenges to measurement, the importance of these sectors emphasizes the value of investing in such research. Because labels are the key determinant of both predictive quality and predictive bias, careful choice can allow us to enjoy the benefits of algorithmic predictions while minimizing their risks.

REFERENCES AND NOTES

1. J. Angwin, J. Larson, S. Mattu, L. Kirchner, "Machine Bias," *ProPublica* (23 May 2016); www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
2. S. Barocas, A. D. Selbst, *Calif. Law Rev.* **104**, 671 (2016).
3. A. Chouldechova, A. Roth, arXiv:1810.08810 [cs.LG] (20 October 2018).
4. A. Datta, M. C. Tschantz, A. Datta, *Proc. Privacy Enhancing Technol.* **2015**, 92–112 (2015).
5. L. Sweeney, *Queue* **11**, 1–19 (2013).
6. M. Kay, C. Matuszek, S. A. Munson, in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (ACM, 2015), pp. 3819–3828.
7. B. F. Klare, M. J. Burge, J. C. Klontz, R. W. Vorder Bruegge, A. K. Jain, *IEEE Trans. Inf. Forensics Security* **7**, 1789–1801 (2012).
8. J. Buolamwini, T. Gebru, in *Proceedings of the Conference on Fairness, Accountability and Transparency* (PMLR, 2018), pp. 77–91.
9. A. Caliskan, J. J. Bryson, A. Narayanan, *Science* **356**, 183–186 (2017).
10. S. Corbett-Davies, S. Goel, arXiv:1808.00023 [cs.CY] (31 July 2018).
11. M. De-Arteaga *et al.*, arXiv:1901.09451 [cs.IR] (27 January 2019).
12. M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, S. Venkatasubramanian, in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2015), pp. 259–268.
13. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, *Q. J. Econ.* **133**, 237–293 (2018).
14. C. S. Hong, A. L. Siegel, T. G. Ferris, *Issue Brief (Commonwealth Fund)* **19**, 1–19 (2014).
15. N. McCall, J. Cromwell, C. Urato, "Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Massachusetts General Hospital and Massachusetts General Physicians Organization (MGH)" (RTI International, 2010).
16. J. Hsu *et al.*, *Health Aff.* **36**, 876–884 (2017).
17. L. Nelson, "Lessons from Medicare's demonstration projects on disease management and care coordination" (Working Paper 2012-01, Congressional Budget Office, 2012).
18. C. Vogeli *et al.*, *J. Gen. Intern. Med.* **22** (suppl. 3), 391–395 (2007).
19. D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, G. Escobar, *Health Aff.* **33**, 1123–1131 (2014).
20. J. Kleinberg, J. Ludwig, S. Mullainathan, Z. Obermeyer, *Am. Econ. Rev.* **105**, 491–495 (2015).
21. G. Hileman, S. Steele, "Accuracy of claims-based risk scoring models" (Society of Actuaries, 2016).
22. J. Kleinberg, S. Mullainathan, M. Raghavan, arXiv:1609.05807 [cs.LG] (19 September 2016).
23. A. Chouldechova, *Big Data* **5**, 153–163 (2017).
24. V. de Groot, H. Beckerman, G. J. Lankhorst, L. M. Bouter, *J. Clin. Epidemiol.* **56**, 221–229 (2003).
25. J. J. Gagne, R. J. Glynn, J. Avorn, R. Levin, S. Schneeweiss, *J. Clin. Epidemiol.* **64**, 749–759 (2011).
26. A. K. Parekh, M. B. Barton, *JAMA* **303**, 1303–1304 (2010).
27. D. Ettehad *et al.*, *Lancet* **387**, 957–967 (2016).
28. K.-T. Khaw *et al.*, *BMJ* **322**, 15 (2001).
29. K. Fiscella, P. Franks, M. R. Gold, C. M. Clancy, *JAMA* **283**, 2579–2584 (2000).
30. N. E. Adler, K. Newman, *Health Aff.* **21**, 60–76 (2002).
31. N. E. Adler, W. T. Boyce, M. A. Chesney, S. Folkman, S. L. Syme, *JAMA* **269**, 3140–3145 (1993).
32. M. Alsan, O. Garrick, G. C. Graziani, "Does diversity matter for health? Experimental evidence from Oakland" (National Bureau of Economic Research, 2018).
33. K. Armstrong, K. L. Ravenell, S. McMurphy, M. Putt, *Am. J. Public Health* **97**, 1283–1289 (2007).
34. M. Alsan, M. Wanamaker, *Q. J. Econ.* **133**, 407–455 (2018).
35. M. van Ryn, J. Burke, *Soc. Sci. Med.* **50**, 813–828 (2000).
36. K. M. Hoffman, S. Trawalter, J. R. Axt, M. N. Oliver, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4296–4301 (2016).
37. J. J. Escarce, F. W. Puffer, in *Racial and Ethnic Differences in the Health of Older Americans* (National Academies Press, 1997), chap. 6; www.ncbi.nlm.nih.gov/books/NBK109841/.
38. S. Passi, S. Barocas, arXiv:1901.02547 [cs.CY] (8 January 2019).
39. S. Mullainathan, Z. Obermeyer, *Am. Econ. Rev.* **107**, 476–480 (2017).
40. K. E. Joynt Maddox *et al.*, *Health Serv. Res.* **54**, 327–336 (2019).
41. K. E. Joynt Maddox, M. Reidhead, A. C. Qi, D. R. Nerenz, *JAMA Intern. Med.* **179**, 769–776 (2019).
42. K. Lum, W. Isaac, *Significance* **13**, 14–19 (2016).
43. I. Ajunwa, "The Paradox of Automation as Anti-Bias Intervention," available at SSRN (2016); https://ssrn.com/abstract=2746078.
44. S. DellaVigna, M. Gentzkow, "Uniform pricing in US retail chains" (National Bureau of Economic Research, 2017).
45. C. A. Gomez-Uribe, N. Hunt, *ACM Trans. Manag. Inf. Syst.* **6**, 13 (2016).

ACKNOWLEDGMENTS

We thank S. Lakhtakia, Z. Li, K. Lin, and R. Mahadeshwar for research assistance and D. Buefort and E. Maher for data science expertise. **Funding:** This work was supported by a grant from the National Institute for Health Care Management Foundation. **Author contributions:** Z.O. and S.M. designed the study, obtained funding, and conducted the analyses. All authors contributed to reviewing findings and writing the manuscript. **Competing interests:** The analysis was completely independent: None of the authors had any contact with the algorithm's manufacturer until after it was complete. No authors received compensation, in any form, from the manufacturer or have any commercial interests in the manufacturer or competing entities or products. There were no confidentiality agreements that limited reporting of the work or its results, no material transfer agreements, no oversight in the preparation of this article (besides ethical oversight from the approving IRB, which was based at a non-profit academic health system), and no formal relationship of any kind between any of the authors and the manufacturer. **Data and materials availability:** Because the data used in this analysis are protected health information, they cannot be made publicly available. We provide instead a synthetic dataset (using the R package synthpop) and all code necessary to reproduce our analyses at https://gitlab.com/labsysmed/dissecting-bias.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/447/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S4
References (*46*–*51*)

8 March 2019; accepted 4 October 2019
10.1126/science.aax2342

NEURODEVELOPMENT

Spatiotemporal expansion of primary progenitor zones in the developing human cerebellum

Parthiv Haldipur[1], Kimberly A. Aldinger[1], Silvia Bernardo[2], Mei Deng[3], Andrew E. Timms[1], Lynne M. Overman[4], Conrad Winter[1], Steven N. Lisgo[4], Ferechte Razavi[5], Evelina Silvestri[6], Lucia Manganaro[2], Homa Adle-Biasette[7,8], Fabien Guilmiot[9], Rosa Russo[10], Debora Kidron[11], Patrick R. Hof[12], Dianne Gerrelli[13], Susan J. Lindsay[4], William B. Dobyns[1,3], Ian A. Glass[1,3], Paula Alexandre[13], Kathleen J. Millen[1,3]*

We present histological and molecular analyses of the developing human cerebellum from 30 days after conception to 9 months after birth. Differences in developmental patterns between humans and mice include spatiotemporal expansion of both ventricular and rhombic lip primary progenitor zones to include subventricular zones containing basal progenitors. The human rhombic lip persists longer through cerebellar development than in the mouse and undergoes morphological changes to form a progenitor pool in the posterior lobule, which is not seen in other organisms, not even in the nonhuman primate the macaque. Disruptions in human rhombic lip development are associated with posterior cerebellar vermis hypoplasia and Dandy-Walker malformation. The presence of these species-specific neural progenitor populations refines our insight into human cerebellar developmental disorders.

Human cerebellar birth defects are common and often cause motor and cognitive disabilities, yet most knowledge of cerebellar development comes from mice. The mouse cerebellum shares many features of lamination, circuitry, neuronal morphology, and foliation with humans. However, compared to the mouse cerebellum, the human cerebellum contains 80% of all brain neurons and has a 750-fold larger surface area, increased neuronal numbers, altered neuronal subtype ratios, and increased folial complexity (*1–3*).

Human cerebellar development begins by 30 days after conception and is complete by the end of postnatal year 2 (*4*, *5*). Mouse cerebellar development is nearly complete by postnatal day 15 (P15), after just 19 gestational days. Although human development is protracted, this alone is unlikely to explain species differences. Cerebral cortical neurogenesis differs between these species, with humans having large numbers of outer radial glia in an outer subventricular zone that drive human cortical expansion and gyrification (*6*, *7*). Because neuronal numbers in the cerebellum and cerebral cortex scale across evolution (*8*), we expected divergent cerebellar neurogenesis programs across species.

Studies of human cerebellar development began before publication of photographic plates (*9*). Limited histological data are available from 10 postconception weeks (PCW) through late gestation (*4*, *5*, *10–12*). Developmental neuroimaging atlases have been compiled from in utero magnetic resonance imaging studies from gestational weeks (GW) 20 to 24 (i.e., 18 to 22 PCW) but are of limited resolution (*13*). Major gaps in available human data correspond to essential cerebellar developmental epochs defined in model vertebrates. Here, we analyze human cerebellar development from 30 days after conception to 9 months after birth (Fig. 1 and table S1), define the timing of developmental events, and provide insight into cellular and

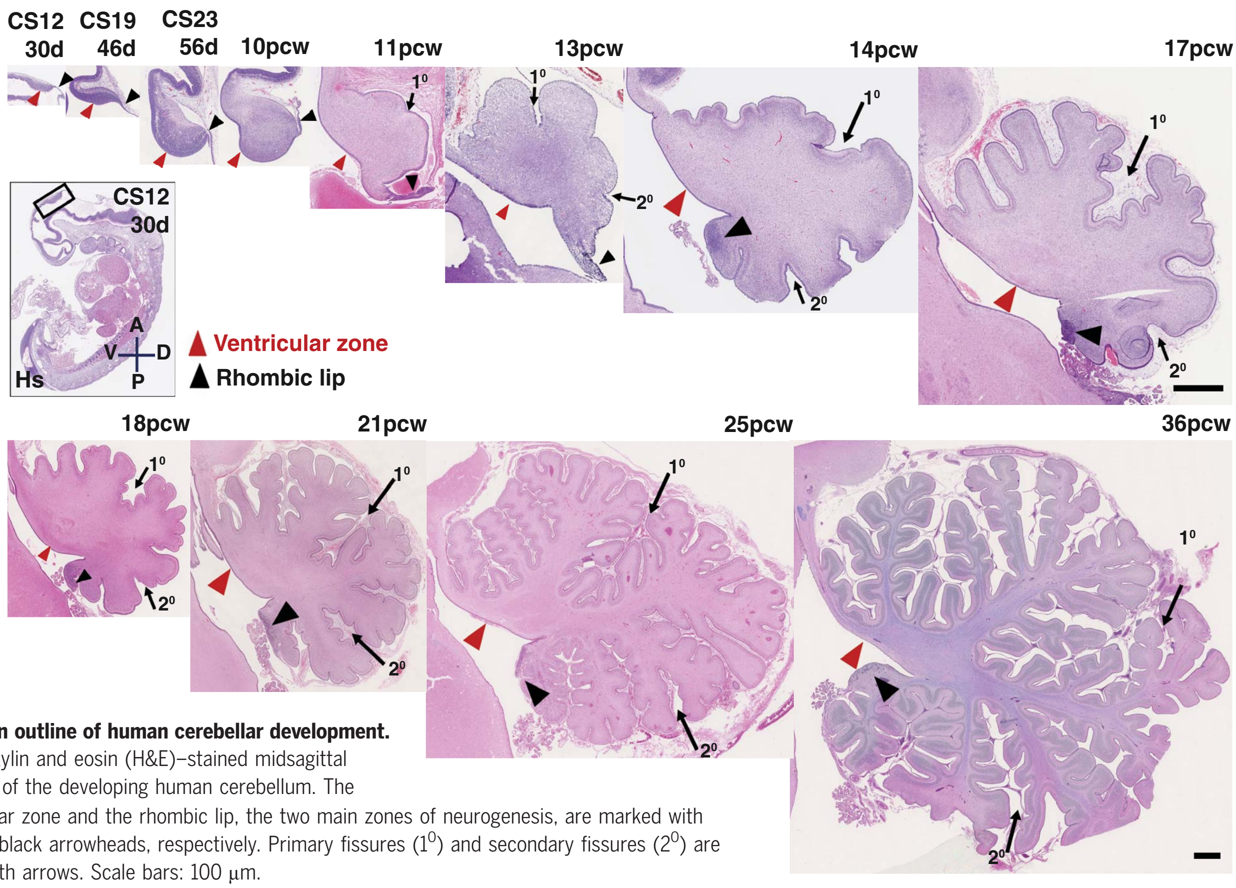


Fig. 1. An outline of human cerebellar development. Hematoxylin and eosin (H&E)–stained midsagittal sections of the developing human cerebellum. The ventricular zone and the rhombic lip, the two main zones of neurogenesis, are marked with red and black arrowheads, respectively. Primary fissures (1^0) and secondary fissures (2^0) are noted with arrows. Scale bars: 100 μm.

molecular programs driving human cerebellar development.

Spatiotemporal expansion of human cerebellar progenitor zones

We surveyed morphology from human and mouse cerebellar sagittal vermis sections. Humans and mice both have two primary zones of neurogenesis: the ventricular zone (VZ) and the rhombic lip (RL) (Fig. 1 and fig. S1A). The VZ gives rise to all γ-aminobutyric acid–releasing (GABAergic) populations, including Purkinje cells. The RL gives rise to all cerebellar glutamatergic neurons. Cerebellar nuclei neurons are generated first, followed by granule cell progenitors of the external granule layer, which proliferate, differentiate, and migrate to become granule neurons of the internal granule layer. Unipolar brush cells arise last (*14*). The human cerebellar anlage between Carnegie stage 12 (CS12) and CS23 (30 to 56 days) resembles the mouse cerebellar anlage in size and shape, from embryonic day 10.5 (E10.5) to E17.5 (Fig. 1 and fig. S1A). However, the mouse VZ thins between E10.5 and E15.5 (fig. S1, B to G), and the RL disappears by birth (E19/P0) (fig. S1A). The human embryonic RL is small, but the human VZ thickens through 10 PCW. After 10 PCW, the VZ thins and the RL expands into an elongated tail-like structure. Between 11 and 13 PCW, the elongated RL thickens, continuing to trail from the growing posterior vermis. Between 13 and 14 PCW, the RL incorporates into the

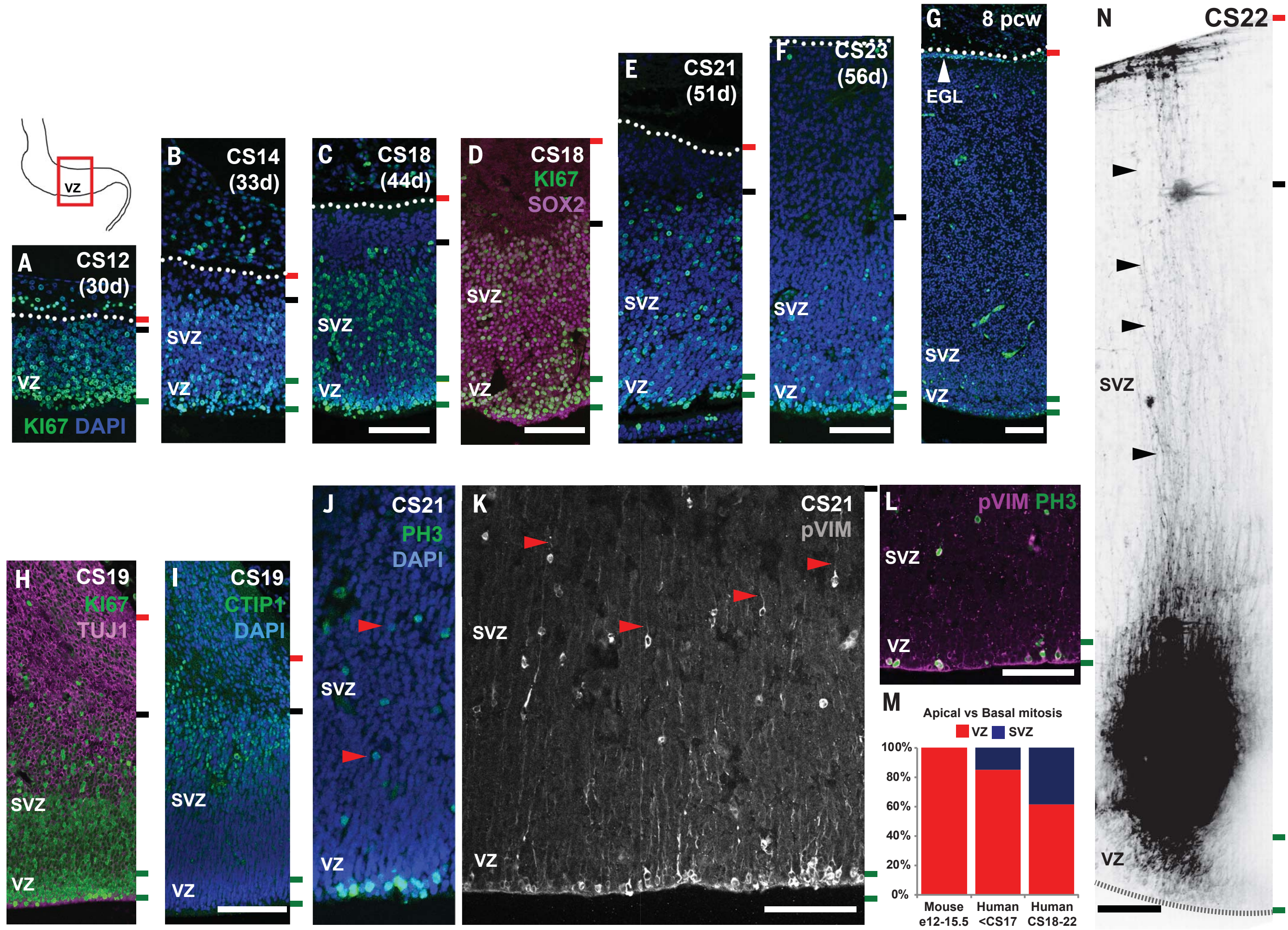


Fig. 2. The human cerebellar VZ is expanded into a SVZ. Midsagittal sections of the human embryonic cerebellum stained with KI67 (**A** to **G**) and SOX2 (D), reveal VZ expansion. (**H**) β-III Tubulin (TUJ1) and (**I**) CTIP1 expression suggest that neuronal differentiation takes place in the SVZ beginning around CS19. (G) External granule layer (EGL) first appears at 8 PCW. (**J** to **L**) Mitosis and mitotic radial glia are observed in both the VZ and SVZ as evidenced by phosphohistone H3 (PH3) (J) and phospho-vimentin (pVIM) [(K) and (L)] expression. (**M**) A significant increase in the proportion of cerebellar basal progenitors [(J) and (K), red arrowheads] is seen between the human and mouse cerebellum, and between CS18 and CS23 (Chi-square, df: 132.5, 2; $P < 0.0001$). (**N**) DiO labeling of VZ and SVZ progenitors at CS22 shows radial glial fibers traversing the thickness of the cerebellum. Sections were counterstained using 4′,6-diamidino-2-phenylindole (DAPI). The boundaries of the VZ, SVZ, and pia are marked with green and black bars along the right side of the images and with dotted lines within the images, respectively. Scale bars: 100 μm.

[1]Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA, USA. [2]Departments of Experimental Medicine and Radiological Sciences, Sapienza University of Rome, Rome, Italy. [3]Department of Pediatrics, University of Washington, Seattle, WA, USA. [4]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [5]Hôpital Necker-Enfants Malades, Assistance Publique Hôpitaux de Paris, Paris, France. [6]Surgical Pathology Unit, San Camillo Forlanini Hospital, Rome, Italy. [7]Department of Pathology, Hôpital Lariboisière, Assistance Publique Hôpitaux de Paris, Paris, France. [8]NeuroDiderot, INSERM, Université de Paris, Paris, France. [9]Hôpital Robert-Debré, INSERM UMR 1141, Paris, France. [10]Department of Pathology, University Medical Hospital, Salerno, Italy. [11]Department of Pathology, Meir Medical Center, Kfar Saba and Sackler School of Medicine, Tel Aviv University, Israel. [12]Nash Family Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [13]University College London Great Ormond Street Institute of Child Health, London, UK.
***Corresponding author. Email: kathleen.millen@seattlechildrens.org**

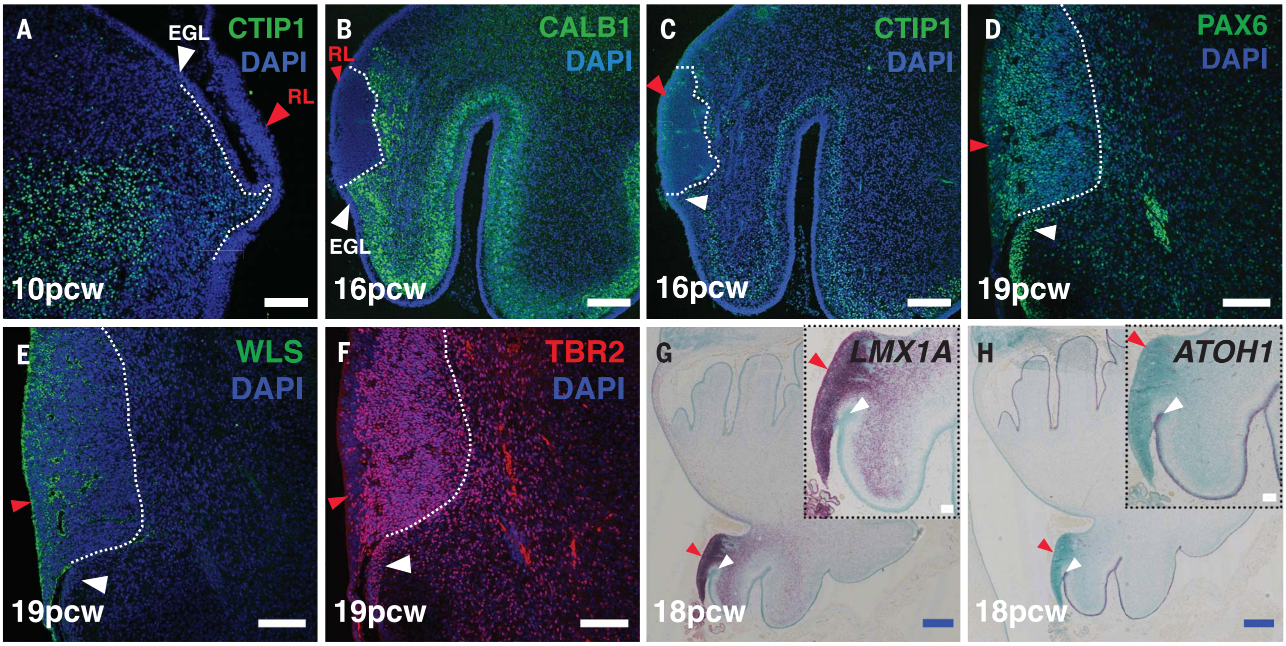


Fig. 3. The human cerebellar rhombic lip expresses classic markers. VZ-born Purkinje cells expressing CTIP1 (**A** and **C**) and calbindin (**B**) migrate around the RL (red arrowhead). The RL expresses classic markers such as (**D**) PAX6, (**E**) WLS, (**F**) TBR2, and (**G**) *LMX1A*. (**H**) *ATOH1* is expressed by cells exiting the RL into the external granule layer (white arrowhead). Sections were counterstained with DAPI [(A) to (F)] and fast green [(G) and (H)]. Scale bars: 100 μm (white) and 500 μm (blue).

posterior lobule where it forms a densely packed pool of cells evident as late as 36 PCW (Fig. 1).

The mouse external granule layer is evident by E12.5, with granule cell progenitor proliferation driving postnatal cerebellar and foliation expansion (*15*, *16*). We detect initial external granule layer (EGL) formation in the human cerebellum at 8 PCW (Fig. 2G), with primary and secondary fissure initiation apparent at 11 and 13 PCW, respectively. Between 17 PCW and birth (~36 PCW) there is an approximately fivefold increase in human cerebellar volume and folial complexity (*17*, *18*). Peak proliferation in the human external granule layer occurs during the period between 26 and 32 PCW (*10*).

The human cerebellum has a SVZ with basal progenitors

At CS12, the human cerebellar VZ resembles the E12.5 mouse VZ which displays a single zone of SOX2$^+$ KI67$^+$ progenitor cells spanning most of the anlage (Fig. 2A and fig. S1, B and F). By CS14, an emerging SVZ is evident (Fig. 2B). By CS18 and CS19, differentiating (TUJ1$^+$CTIP1$^+$) neurons increase cerebellar anlage size (Fig. 2, H and I). Increased differentiation in the outer SVZ diminishes SVZ size between CS21 and CS23 (Fig. 2, E and F). By the end of embryogenesis, at 8 PCW, only a residual VZ remains (Fig. 2G).

The expanded proliferative zone in the embryonic human cerebellum resembles the SVZ in the developing mouse and human cerebral cortex (*19*). The mouse cerebellar VZ does not have a SVZ. Instead, ventricular radial glial progenitors extend processes across the nascent anlage from the ventricular (apical) to the pial (basal) surface and undergo mitosis only at the ventricle (Fig. 2, M and N, and fig. S1, L to Q). In the human cerebellum, mitotic phosphohistone H3$^+$ (PH3$^+$) progenitors are found within the VZ and SVZ, indicating the presence of basal progenitors, which we call cerebellar basal progenitors [Fig. 2, J and K (red arrowheads), and M]. Progenitors in both zones exhibit long radial processes which span the anlage thickness (Fig. 2N) and express mitotic radial glial marker phospho-vimentin (Fig. 2K). We see a significant expansion of cerebellar basal progenitors throughout the SVZ between CS18 and CS23 (Fig. 2M), a time point coinciding with increased differentiation, suggesting they could function as additional neurogenic progenitors.

The human cerebellar RL is long-lived and compartmentalized

Species differences in progenitor zone development are not restricted to the cerebellar VZ. We also identified differences in RL morphology, finding substructure in humans, including a split into ventricular (RL^{VZ}) and subventricular zones (RL^{SVZ}), as well as internalization (Fig. 3 and fig. S3, A to G). The mouse RL is a proliferative, transient, dorsal stem cell zone, present between E12.5 and E17.5 and composed entirely of KI67$^+$ and SOX2$^+$ stem cells lacking morphological compartmentalization (fig. S2, A to E). Although it is 5 to 8 cell layers thick, progenitor mitosis (PH3$^+$ and phospho-vimentin$^+$) is confined to the single layer of cells lining the ventricle (fig. S2, G to L). In contrast, the human RL and its remnants are seen throughout gestation and display a more complex proliferation profile (figs. S3, A to G, and S5C).

The human RL excludes VZ-derived GABAergic neurons, expresses classic RL markers, and is proliferative even when embedded in the posterior-most lobule of the cerebellum (Figs. 3, A to H, and 4B; and fig. S3, A to K). The human embryonic RL, similar to that of the mouse, consists of KI67$^+$ and SOX2$^+$ cells (Fig. 4, B and C, CS18; and fig. S2, A to F). However, after 10 PCW, the human RL splits into a SOX2$^+$ and KI67–rich RL^{VZ} (Fig. 4, B and C, red asterisk) and a KI67-rich, SOX2-sparse RL^{SVZ} (Fig. 4, B and C, yellow asterisk). The two RL progenitor zones are separated by a vasculature bed, discernible by 11 PCW (fig. S5, A and B, arrows). Cells in the RL^{SVZ} apparently migrate into the external granule layer (Fig. 4, B and C; and fig. S3, H and I, white arrowheads).

Radial glial mitosis in the early human RL is confined to cells lining the ventricle (Fig. 4, D and E, CS18), much like in the mouse (fig. S2, G to L). However, after the split of the RL, mitotic progenitors with radial glia–like morphology (PH3$^+$, phospho-vimentin$^+$) are detected in both the RL^{VZ} and RL^{SVZ} (Fig. 4, D and E, 13 PCW). DiI and lentiviral labeling reveal diverse morphologies of RL basal progenitors including uni-, bi-, and multipolar cells with radially and tangentially oriented processes (Fig. 4, A and F).

Although the mouse RL lacks structural compartmentalization, it is molecularly compartmentalized. An interior Wntless (WLS$^+$) and

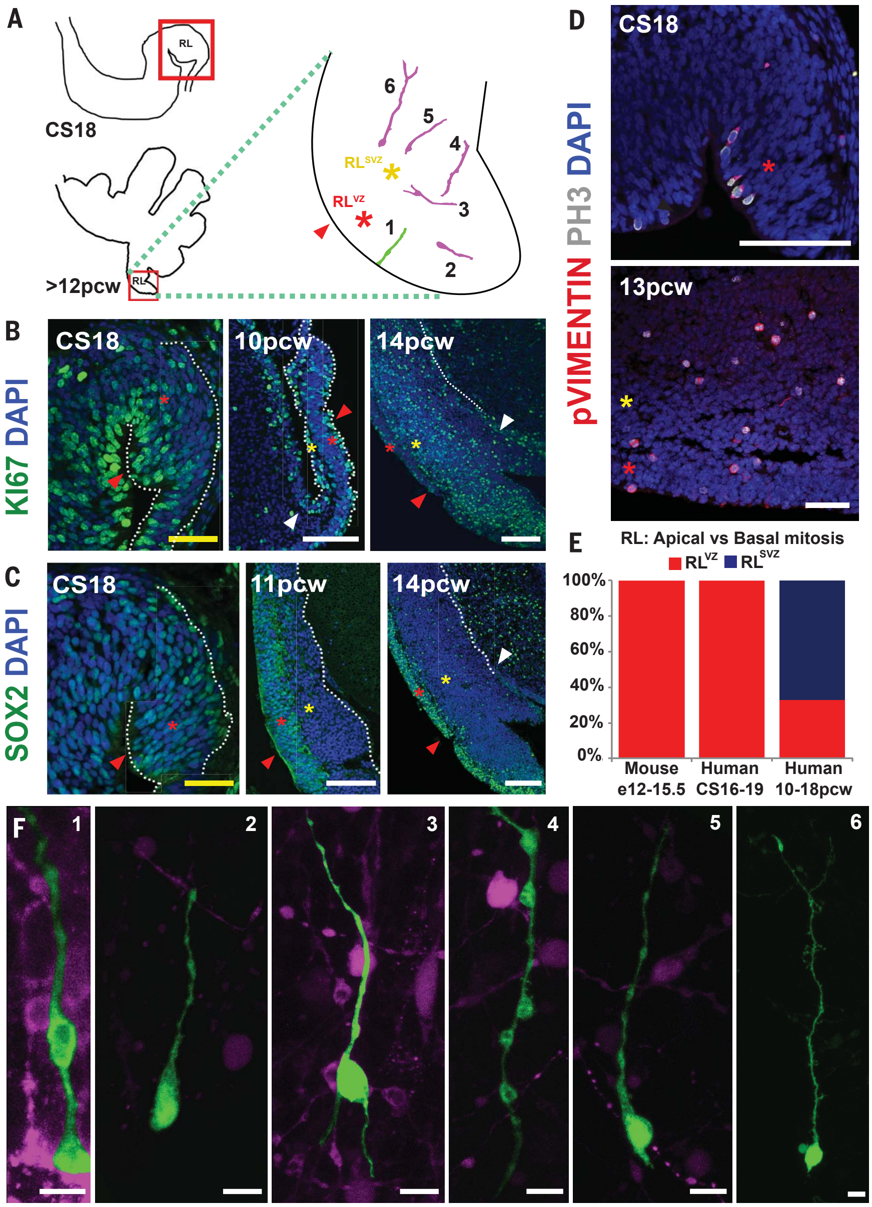


Fig. 4. The human cerebellar RL is compartmentalized into ventricular and subventricular zones. (**A**) Illustration of the cerebellar regions studied in this figure. The RL (red arrowhead) is expanded into ventricular (RL^{VZ}, red asterisk) and subventricular zones (RL^{SVZ}, yellow asterisk). (**B**) KI67 expression reveals extensive proliferation in the RL. (**C**) The RL is compartmentalized into a SOX2-rich RL^{VZ} and SOX2-sparse RL^{SVZ}. (**D**) Phosphohistone-H3 and phospho-vimentin expression indicate the presence of mitotic ventricular and subventricular-basal progenitors. (**E**) The proportion of basal progenitors increases significantly after the splitting of the RL (Chi-square, df: 137.8, 2; $P < 0.0001$). (**F**) DiI and lentiviral labeling of an organotypic slice of the human cerebellum reveal diverse morphologies of RL ventricular and basal progenitors. Scale bars: 100 μm (white) and 50 μm (yellow).

$LMX1A^+$ compartment is continuous with the VZ and an exterior compartment links to the external granule layer ($ATOH1^+$). A gradient of $PAX6^+$ expression exists across the mouse RL, with strongest expression in the exterior compartment. The core of the mouse RL is composed of proliferating $LMX1A^+$ progenitors destined to become posterior vermis granule cell progenitors and unipolar brush cells. Early specified and differentiating unipolar brush cells in the core also express TBR2 (*20*, *21*).

In humans, WLS expression is also largely restricted to RL^{VZ} cells, although scattered expression was seen in RL^{SVZ} cells (Fig. 3E and fig. S4, A to C). *LMX1A* is expressed throughout the embryonic RL (fig. S4D), in both the RL^{VZ} and RL^{SVZ} at later stages (Fig. 3G and fig. S4, E and F). *LMX1A* is also expressed in RL-derived cerebellar nuclear and unipolar brush cell populations streaming into the cerebellar core, as well as the choroid plexus epithelium (fig. S4E). A sharp boundary between *LMX1A* and *ATOH1* in RL exiting granule cell progenitors define the anterior limit of the RL (Fig. 3, G and H, and fig. S4, D to J). The posterior limit of the human RL is defined by $LMX1A^+$ $MKI67^-$ choroid plexus cells (fig. S4, D, G, and H). PAX6 expression is predominant in the RL^{VZ}, although there is also extensive expression in the RL^{SVZ}, with up-regulation in nascent external granule layer cells streaming from the RL^{SVZ} (Fig. 3D and fig. S4, K to M). TBR2 is expressed throughout the RL^{SVZ} with a few scattered, presumably nascent unipolar brush cells, in both the external and internal granule layer (Fig. 3F and fig. S4, N to P).

Human RL progenitors share some similarities with mice

To provide an unbiased analysis of the molecular programs encoded by human RL progenitors, we profiled the human RL transcriptome using bulk sequencing from laser-capture microdissected RL^{VZ} and RL^{SVZ} (data S1 to S4). Principal component analysis indicated age as the first principal component, explaining 56% of the variance (Fig. 5A). Differential expression analysis between RL^{VZ} and RL^{SVZ} identified alterations in 622 genes ($\log_2$ fold change > 1.5 and Benjamini-Hochberg adjusted $P < 0.05$). The 374 genes up-regulated in RL^{VZ} included *CRYAB*, *SOX2*, and *WLS*, and the 248 genes up-regulated in the RL^{SVZ} included *EOMES* (Fig. 5B and data S2). We evaluated pathway enrichment of the up-regulated genes and found RL^{VZ} genes were enriched in HIPPO and WNT signaling (Fig. 5C and data S3). RL^{SVZ} genes were enriched in axon guidance and synaptic vesicle cycling (data S4). Several known mouse marker genes for RL and early RL derivates showed similar expression profiles in our samples (*22*) (fig. S3L). We next compared genes with significant differential expression between the RL^{VZ} and RL^{SVZ} to mouse gene sets identified in the RL and RL-related cells in recent single cell analyses of the developing mouse cerebellum (*23*, *24*). Whereas roof plate–like stem cell genes and a subset of mouse RL genes were differentially expressed in the RL^{VZ}, genes expressed in mouse RL, unipolar brush cell, and granule cell progenitors were seen throughout the RL. Unipolar brush cell genes were highly expressed in the RL^{SVZ}, indicating that it is likely a reservoir for nascent human unipolar brush cells, in addition to other glutamatergic lineages (Fig. 5, D to G, and fig. S3M). Our analyses validate our hypothesis that the spatiotemporally expanded structure in the posterior vermis is indeed the RL, and similarities in gene expression patterns exist between the mouse and human RL notwithstanding differences in structure.

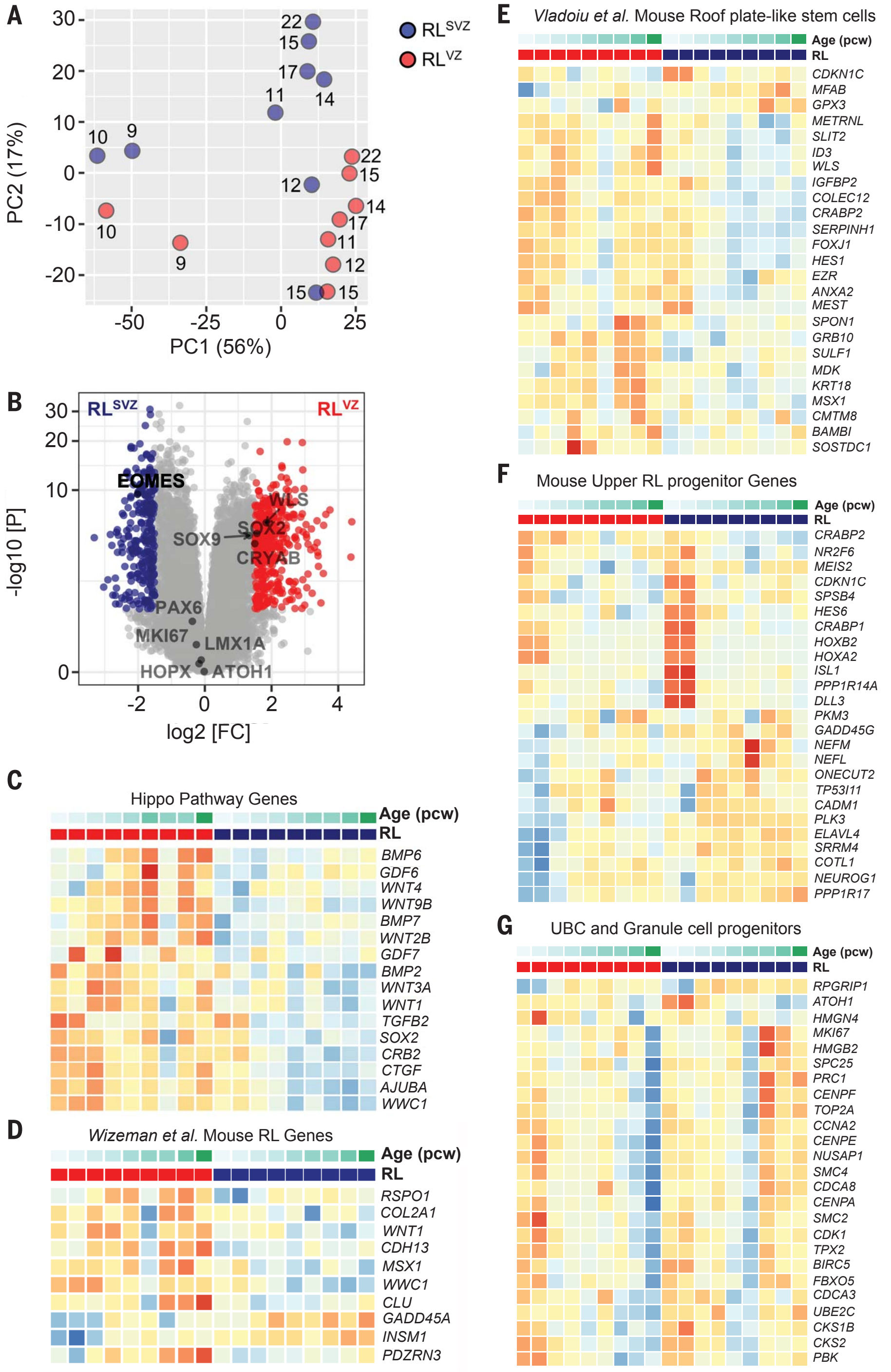


Fig. 5. RNA-seq of human RL compartments. (**A**) Principal component (PC) analysis indicates that the largest source of variation among the RNA sequencing (RNA-seq) samples was age, accounting for 56% of the variance in the data. Samples microdissected from RL^{SVZ} are blue and those from RL^{VZ} are red. The numbers beside each circle represent sample age (PCW, postconception weeks). (**B**) Volcano plot illustrating differential expression of genes in RL^{VZ} versus RL^{SVZ}. Red and blue dots represent genes expressed significantly higher in RL^{VZ} or RL^{SVZ}, respectively. (**C** to **G**) Heatmaps of gene expression for each human sample. Samples are grouped by RL^{VZ} (red) and RL^{SVZ} (blue), then by ascending age (9 to 22 PCW). Expression of (C) Hippo pathway genes and [(D) to (G)] top genes for RL-related cell clusters identified by single-cell RNA-seq of the mouse cerebellum from E10 to P14 (*23*, *24*).

RL spatiotemporal expansion may be a human-specific feature

The developing human and nonhuman primate cerebral cortices share an expanded SVZ relative to mice (*25*, *26*). Because the brain weight–to–neuron number ratio does not differ significantly among primates, we expected nonhuman primates to share elaborated cerebellar progenitor zones (*27*, *28*). We analyzed midsagittal sections of the developing cerebella of rhesus macaque (*Macaca mulatta*; 164-day gestation). The macaque RL at E48 resembles the embryonic CS23 RL in humans (fig. S6, A and M). However, as foliation initiates, the macaque RL regresses in a manner similar to the mouse RL and unlike humans (fig. S6, B and C). By mid to late gestation (E78 to E133), there is no evidence of RL expansion or morphological compartmentation (fig. S6, D to L). This suggests that spatiotemporal expansion of the RL may be specific to humans.

The internalized RL generates the posterior vermis

Human cerebellar volume increases fivefold between 22 PCW and birth, and it becomes highly foliated during the third trimester (24 to 40 GW) (*17*, *18*). Granule cell progenitor proliferation peaks during this period, accompanied by increased external granule layer thickness (*5*, *10*). In mice, cerebellar growth and foliation are driven by granule cell progenitor proliferation between P1 and P14, with deficient proliferation causing external granule layer thinning, reduced foliation, and reduced cerebellar volume (*15*, *16*). Disproportionate reduction in posterior vermis volume is a feature of many human cerebellar birth defects, including Dandy-Walker malformation and cerebellar vermis hypoplasia (*29*). Mouse models have indicated that RL disruption in the form of precocious differentiation or aberrant migration of progenitors is central to posterior vermis hypoplasia (*30*–*32*). In mice and humans, 10 cardinal lobules are grouped together into anterior, central, and posterior lobes of the vermis defined by the primary and secondary fissure. In humans, all 10 lobules are first identifiable between 14 and 18 PCW. Development progresses from anterior to posterior, with the increase in posterior vermis size and folial complexity relative to the anterior cerebellum beginning only after 17 PCW (Fig. 6A). The RL and its vestiges are detectable in human cerebella until birth. The longevity of the internalized RL thus correlates with growth and foliation of the human posterior vermis, a cerebellar region associated with human cognition (*33*).

To determine whether RL abnormalities specifically contribute to posterior vermis hypoplasia in Dandy-Walker malformation and cerebellar vermis hypoplasia, we assessed 16 archival samples aged 17 to 30 PCW (table S1). All show delay or failure of posterior vermis

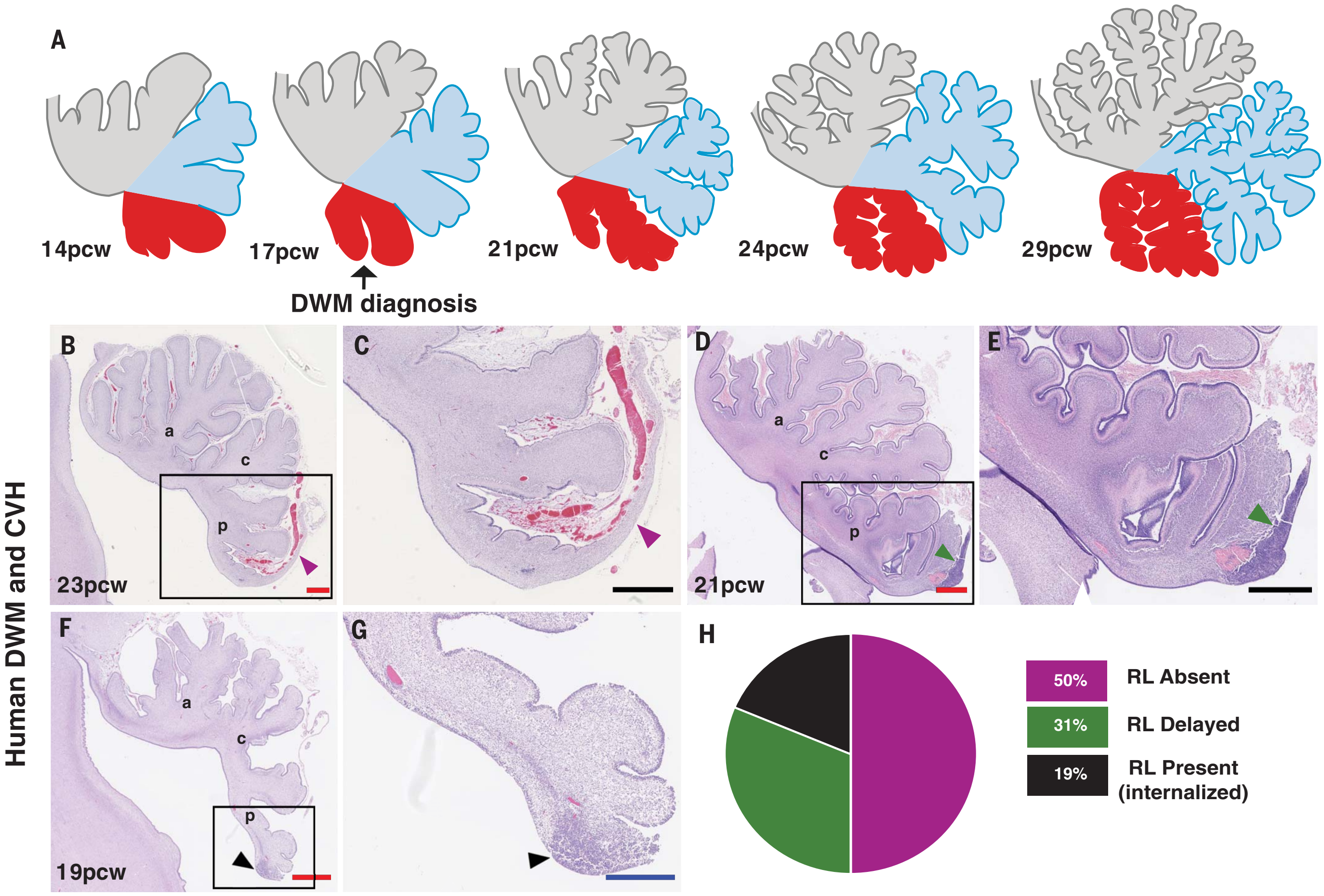


Fig. 6. Internalized RL may be a feature specific to human cerebellar development. (**A**) Model for human cerebellar development indicates that growth of the posterior vermis correlates with spatiotemporal RL expansion. (**B** to **G**) Analysis of H&E-stained sagittal sections of the human cerebellum from cases diagnosed with Dandy-Walker malformation (DWM) and cerebellar vermis hypoplasia (CVH) indicates that RL is absent in 50% of cases [(B) and (C)] and is severely diminished in the remaining cases [(D) to (G)]. The anterior (a), central (c), and posterior (p) lobes are indicated in gray, blue, and red, respectively. Colored arrowheads correspond to the colored features in (H). (**H**) Pie chart representing the absence of RL in 50% of tested samples. Scale bars: 100 μm (black), 0.5 mm (blue), and 1 mm (red).

growth. Most show only a partially formed posterior-most lobule (*31*). Although overall decreased external granule layer proliferation may contribute to this phenotype, peak external granule layer proliferation only begins around 26 PCW (26 to 32 PCW), and hypoplasia of the posterior vermis is evident earlier (Fig. 6, B to H, and fig. S7, A to P) (*30*, *31*, *34*). Indeed, Dandy-Walker malformation and cerebellar vermis hypoplasia are routinely diagnosed in utero around 17 PCW (*35*). The RL is absent in half of our cases and delayed (not internalized) or reduced in size and cellularity in the others (Fig. 6H). Among cerebella lacking a RL, ~62% are older than 21 PCW, when the human RL is normally still present. We conclude that whereas the early RL (<13 PCW) contributes granule cell progenitors to the external granule layer which later proliferate to drive expansion of the anterior cerebellum, the older, internalized RL (>14 PCW) generates granule cell progenitors required to fully elaborate the posterior vermis during mid and late gestation.

Preterm brain injury is associated with cerebellar abnormalities (*36*, *37*). We find RL^{VZ} and RL^{SVZ} are separated by a vascular bed beginning at 11 PCW, when the RL itself is elongated and perhaps more vulnerable to insult. Vascular insults causing RL structural damage from 13 to 14 PCW onward likely contribute to these neurodevelopmental abnormalities. Thus, although mouse studies were essential to spotlight a role for the RL in human Dandy-Walker malformation, the underlying pathological mechanisms can likely never be fully modeled in mice that lack complex RL anatomy. Similarly, our discovery of previously undescribed progenitor populations and the longevity of the RL in the human cerebellum suggest that mouse models of the cells of origin for some groups of medulloblastoma, a cerebellar tumor, may also be inadequate (*38*). Our studies underline the urgency of further comparative cellular and molecular analyses of human and mouse cerebellar development to better define the value and limitations of mouse genetic models of human neurodevelopmental disorders.

REFERENCES AND NOTES

1. W. Lange, *Cell Tissue Res.* **157**, 115–124 (1975).
2. D. C. Van Essen, *Ann. N. Y. Acad. Sci.* **978**, 468–479 (2002).
3. O. Larsell, J. Jansen, *The Comparative Anatomy and Histology of the Cerebellum. The Human Cerebellum, Cerebellar Connections, and Cerebellar Cortex* (Univ. of Minnesota Press, 1972).
4. P. Haldipur *et al.*, *PLOS ONE* **6**, e23449 (2011).
5. P. Rakic, R. L. Sidman, *J. Comp. Neurol.* **139**, 473–500 (1970).
6. Z. Molnár, A. Pollen, *Development* **141**, 11–16 (2014).
7. T. J. Nowakowski, A. A. Pollen, C. Sandoval-Espinosa, A. R. Kriegstein, *Neuron* **91**, 1219–1227 (2016).
8. S. Herculano-Houzel, *Front. Neuroanat.* **4**, 12 (2010).
9. O. Larsell, W. A. Stotler, *Anat. Rec.* **97**, 352 (1947).
10. H. Ábrahám, T. Tornóczky, G. Kosztolányi, L. Seress, *Int. J. Dev. Neurosci.* **19**, 53–62 (2001).
11. P. Haldipur *et al.*, *Stem Cells Dev.* **21**, 1059–1068 (2012).
12. N. Zecevic, P. Rakic, *J. Comp. Neurol.* **167**, 27–47 (1976).
13. P. A. Habas *et al.*, *Neuroimage* **53**, 460–470 (2010).
14. K. Leto *et al.*, *Cerebellum* **15**, 789–828 (2016).
15. J. D. Corrales, G. L. Rocco, S. Blaess, Q. Guo, A. L. Joyner, *Development* **131**, 5581–5590 (2004).

16. N. Dahmane, A. Ruiz i Altaba, *Development* **126**, 3089–3100 (1999).
17. C. Limperopoulos *et al*., *Pediatrics* **115**, 688–695 (2005).
18. J. J. Volpe, *J. Child Neurol.* **24**, 1085–1104 (2009).
19. A. A. Pollen *et al*., *Cell* **163**, 55–67 (2015).
20. V. V. Chizhikov *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10725–10730 (2010).
21. J. Yeung *et al*., *J. Neurosci.* **34**, 12527–12537 (2014).
22. D. Morales, M. E. Hatten, *J. Neurosci.* **26**, 12226–12236 (2006).
23. M. C. Vladoiu *et al*., *Nature* **572**, 67–73 (2019).
24. J. W. Wizeman, Q. Guo, E. M. Wilion, J. Y. Li, *eLife* **8**, e42388 (2019).
25. M. Florio, W. B. Huttner, *Development* **141**, 2182–2194 (2014).
26. J. H. Lui, D. V. Hansen, A. R. Kriegstein, *Cell* **146**, 18–36 (2011).
27. F. A. Azevedo *et al*., *J. Comp. Neurol.* **513**, 532–541 (2009).
28. S. Herculano-Houzel, *Front. Hum. Neurosci.* **3**, 31 (2009).
29. A. J. Barkovich, K. J. Millen, W. B. Dobyns, *Brain* **132**, 3199–3230 (2009).
30. K. A. Aldinger *et al*., *Nat. Genet.* **41**, 1037–1042 (2009).
31. P. Haldipur *et al*., *eLife* **6**, e20898 (2017).
32. P. Haldipur *et al*., *eLife* **3**, e03962 (2014).
33. J. D. Schmahmann, X. Guell, C. J. Stoodley, M. A. Halko, *Annu. Rev. Neurosci.* **42**, 337–364 (2019).
34. S. Bernardo *et al*., *Prenat. Diagn.* **35**, 1358–1364 (2015).
35. A. Poretti, E. Boltshauser, T. A. G. M. Huisman, *Cerebellum* **15**, 5–9 (2016).
36. K. A. Aldinger *et al*., *Am. J. Hum. Genet.* **105**, 606–615 (2019).
37. A. Pichiecchio *et al*., *Eur. J. Paediatr. Neurol.* **20**, 188–191 (2016).
38. R. Azzarelli, B. D. Simons, A. Philpott, *Development* **145**, dev162693 (2018).

ACKNOWLEDGMENTS

We thank B. Lopez, N. Moreno, M. Crosier, Y. Cheng, J. Dobor (HDBR), D. O'Day (UW), A. Sjoboen, J. Millman, D. Dang, D. Dubocanin (SCRI), B. Wicinski (Hof laboratory), A. Thrasher (UCL), A. Duque and L. Selemon (MacBrainResource), and NIHR GOSH Biomedical Research Centre for providing resources and technical help. We thank D. Price (Edinburgh) for recommending HDBR, and S. Tole (TIFR, India) for feedback on the manuscript. **Funding:** This work was supported by NIH-R01-NS080390 and R01-NS095733 to K.J.M. and R01-N5050375 to W.B.D. P.H. was awarded EMBO fellowship ATSF-431-2016, Burroughs-Wellcome Fund 1018771, Company of Biologists Fellowship DEVTF190393, and a National Ataxia Foundation Young Investigator Research Grant. P.A. was awarded the Newlife Charity for Disabled Children Start-Up Grant SG/17-18/05. **Author contributions:** Conceptualization: P.H., K.J.M.; Methodology: P.H., K.A.A., P.A., K.J.M.; Software: K.A.A., A.E.T., P.A.; Validation: P.H., L.M.O.; Formal analysis: P.H., P.A., K.A.A., K.J.M.; Investigation: P.H., M.D., C.W., L.M.O., P.A.; Resources: S.B., S.N.L., I.A.G., D.G., S.J.L., D.K., L.M., R.R., H.A.-B., F.R., E.S.F.G.; Data Curation: P.H., K.A.A., A.E.T.; Writing – original draft preparation: P.H., P.A., K.A.A., K.J.M.; Writing – review and editing: P.H., K.A.A., P.R.H., P.A., K.J.M.; Visualization: P.H., K.A.A., P.A., W.B.D., K.J.M.; Supervision: P.H., K.J.M.; Project administration: P.H., K.J.M.; Funding acquisition: K.J.M. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Human material was provided by the Joint MRC/Wellcome (MR/R006237/1) Human Developmental Biology Resource (www.hdbr.org) and the Birth Defects Research Laboratory (NIH-R24-HD000836 to I.A.G.) and covered by a material transfer agreement between SCRI and HDBR/BDRL, but samples may be requested directly from the HDBR/BDRL. Macaque images were provided by MacBrainResource (macbrainresource.org; NIMH-R01-MH113257 to A.D. and L.S.). Most sequence data was deposited into dbGaP, the Database of Genotypes and Phenotypes under accession number phs001901.v1.p1, with remaining data available upon request with data use agreement.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/454/suppl/DC1
Materials and Methods
Figs. S1 to S7
Table S1
References (*39*–*45*)
Data S1 to S4

View/request a protocol for this paper from *Bio-protocol*.

18 April 2019; accepted 25 September 2019
10.1126/science.aax7526

INNATE IMMUNITY

Palmitoylation of NOD1 and NOD2 is required for bacterial sensing

Yan Lu[1,2]*, Yuping Zheng[1]*, Étienne Coyaud[3]*, Chao Zhang[4,5]*, Apiraam Selvabaskaran[6], Yuyun Yu[1], Zizhen Xu[1], Xialian Weng[1], Ji Shun Chen[1], Ying Meng[1], Neil Warner[7], Xiawei Cheng[8], Yangyang Liu[9], Bingpeng Yao[10], Hu Hu[9], Zonping Xia[11], Aleixo M. Muise[7], Amira Klip[12], John H. Brumell[7,13], Stephen E. Girardin[6], Songmin Ying[10], Gregory D. Fairn[2]†, Brian Raught[3,14]†, Qiming Sun[8]†, Dante Neculai[1]†

The nucleotide oligomerization domain (NOD)–like receptors 1 and 2 (NOD1/2) are intracellular pattern-recognition proteins that activate immune signaling pathways in response to peptidoglycans associated with microorganisms. Recruitment to bacteria-containing endosomes and other intracellular membranes is required for NOD1/2 signaling, and NOD1/2 mutations that disrupt membrane localization are associated with inflammatory bowel disease and other inflammatory conditions. However, little is known about this recruitment process. We found that NOD1/2 S-palmitoylation is required for membrane recruitment and immune signaling. ZDHHC5 was identified as the palmitoyltransferase responsible for this critical posttranslational modification, and several disease-associated mutations in NOD2 were found to be associated with defective S-palmitoylation. Thus, ZDHHC5-mediated S-palmitoylation of NOD1/2 is critical for their ability to respond to peptidoglycans and to mount an effective immune response.

The cytosolic pattern recognition receptors (PRRs) nucleotide oligomerization domain 1 (NOD1) and NOD2 play crucial roles in host defense and survival, primarily by conferring responsiveness to cytosolic bacterial peptidoglycans [γ-D-glutamyl-*meso*-diaminopimelic acid (iE-DAP) and muramyl dipetide (MDP)] shed by bacteria during infection (*1*, *2*). Dysregulation of NOD1/2 function leads to severe immunologic and inflammatory diseases such as Crohn's disease (CD) and Blau syndrome (*1*, *2*). A NOD2 variant (3020insC frameshift mutation that leads to a truncated NOD2 protein) is implicated in the pathogenesis of CD (*3*, *4*). Although soluble in the cytosol, NOD1/2 associate with the plasma membrane (PM) and endosomal compartments (*5*–*8*) for the surveillance of bacterial cell wall components and promote activation of the nuclear factor κB (NF-κB) and mitogen-activated protein kinase MAPK) signaling pathways from endosomal membranes by means of the RIP2 kinase (*9*). Although fully competent to bind MDP (*10*), membrane association is not observed in the NOD2^{3020insC} variant, suggesting that membrane localization of NOD1/2 is essential for their function (*6*). Lacking recognizable membrane-targeting domains (*11*), NOD1/2 have been suggested to be anchored to membranes indirectly, through cytoskeletal components or membrane-bound proteins (*12*–*16*) or to endosomes by endosomal proteins such as SLC15A3 (*6*). However, these models do not fully explain the rapid redistribution of NODs from plasmalemma to endosomal compartments in response to bacterial invasion (*17*).

Soluble proteins can be targeted to cellular membrane structures by lipidation (*18*). Palmitoylation, unlike farnesylation or geranylgeranylation, does not require a consensus motif in target proteins and is widely implicated in the regulation of protein localization, trafficking, and stability (*18*–*20*). S-palmitoylation, catalyzed by the ZDHHC domain–containing protein acyl-transferases (PATs) (*21*), also plays important roles in immune responses (*22*).

NOD1 and NOD2 are S-palmitoylated

Treatment with 2-bromopalmitate (2BP), an inhibitor of the ZDHHC PATs (*21*, *23*), effected the redistribution of green fluorescent protein (GFP)–NOD1 and GFP-NOD2 from plasmalemmal and endosomal membranes (Fig. 1A and fig. S1A) to the cytosol in RAW264.7, HCT116 (Fig. 1B), and human embryonic kidney (HEK) 293 cells (fig. S1B), mimicking the phenotype reported for the CD mutant protein NOD2^{3020insC} (Fig. 1C and fig. S1C) (*6*). To quantify the effect of 2BP on NOD1/2 PM localization, we used the split superfolder GFP (sfGFP) system combined with dual-color flow cytometry (*24*). Bicistronic plasmids were created that coexpress seven copies of GFP11 tethered to the plasmalemma by the transmembrane domain of the PDGF receptor (HA-TM^{PDGF}-GFP11$_{x7}$) and the GFP1–10 fragment alone (as control) or fused to NOD1 or NOD2 at their N termini. Robust reconstituted GFP fluorescence (GFP$_{comp}$) was observed at the PM of 293 cells (fig. S1, D and E). 2BP treatment reduced the PM localization of NOD1/2 that prevented the formation of functional GFP to levels comparable with that of controls (Fig. 1, D and E). The subcellular distribution of endogenous NOD1 (fig. S1F) in primary mouse bone marrow–derived macrophages (*m*BMDMs) (Fig. 1F) and primary human monocyte–derived macrophages (*h*MDMs) (fig. S1G) was similarly altered by 2BP treatment (Fig. 1G and fig. S1H). 2BP treatment of primary *m*BMDMs (Fig. 1H and fig. S1I) or the colorectal epithelial cell line HCT116 (fig. S1, J and K) also substantially reduced the levels of endogenous NOD1 (fig. S1, L and M) or NOD2 (fig. S1, N and O) associated with the insoluble membrane fraction but did not affect total NOD1/2 protein levels (Fig. 1H and fig. S1, I to K). By contrast, the subcellular distribution of NOD1/2 was not affected by inhibition of farnesyltransferase or geranyltransferase activity by using FTI-276 or GGTI-2133, respectively (fig. S1P).

To examine NOD1/2 palmitoylation, we performed an acyl-biotin exchange (ABE) assay (*25*). Ectopically expressed NOD1/2 were acylated (Fig. 1I), and their acylation levels were reduced by more than 75% upon treatment with 2BP (Fig. 1, J and K). Loss of signal upon omission of hydroxylamine (HAM) treatment demonstrated that NOD1/2 incorporate palmitate through a thioester linkage (Fig. 1I). Metabolic incorporation of the bioorthogonal fatty acid analog 17-octadecynoic acid (17-ODYA) (fig. S1R) was also monitored with click chemistry

[1]Department of Cell Biology, and Department of Pathology Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [2]Keenan Research Centre for Biomedical Sciences, St. Michael's Hospital, Toronto, ON, Canada. [3]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada. [4]Department of Anatomy, School of Medicine, Zhejiang University, China School of Medicine, Zhejiang University, Hangzhou 310058, China. [5]Key Laboratory of Respiratory Disease of Zhejiang Province, Department of Respiratory and Critical Care Medicine, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China. [6]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada. [7]SickKids Inflammatory Bowel Disease Center and Cell Biology Program, Research Institute, Hospital for Sick Children, Toronto, ON, Canada. [8]Department of Biochemistry, and Department of Cardiology of Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China. [9]Department of Pathology, School of Medicine, Zhejiang University, Hangzhou 310058, China. [10]Department of Pharmacology and Key Laboratory of Respiratory Disease of Zhejiang Province, Department of Respiratory and Critical Care Medicine, Second Affiliated Hospital, Institute of Respiratory Diseases, Zhejiang University School of Medicine, Hangzhou 310009, China. [11]Translational Medicine Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China. [12]Cell Biology Program, Research Institute, The Hospital for Sick Children, Peter Gilgan Centre for Research and Learning (PGCRL), Toronto, ON, Canada. [13]Department of Molecular Genetics and Institute of Medical Science, University of Toronto, Toronto, ON, Canada. [14]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada.
*These authors contributed equally to this work.
†Corresponding author. Email: greg.fairn@unityhealth.to (G.D.F.); brian.raught@uhnresearch.ca (B.R.); qmsun@zju.edu.cn (Q.S.); dneculai@zju.edu.cn (D.N.)

conjugation and fluorescent detection (Fig. 1, L and M) (*26*).

2BP impaired downstream signaling triggered by the PRRs, as indicated by reduced NF-κB activity in RAW267.4 and HCT116 cells treated with the iE-DAP analog C12-iE-DAP or MDP (Fig. 1, N and O, and fig. S1, S and T). These results agree with the ability of 2BP to impair the release of cytokines in C12-iE-DAP– and MDP-stimulated *m*BMDMs (fig. S1, U and V) (*6*). Furthermore, treatment with cerulenin (*23*), a drug that inhibits S-palmitoylation, yielded similar levels of NF-κB phospho-p65 and phospho-p38 inhibition (fig. S1, W and X) in RAW264.7 cells. Together, these data indicate

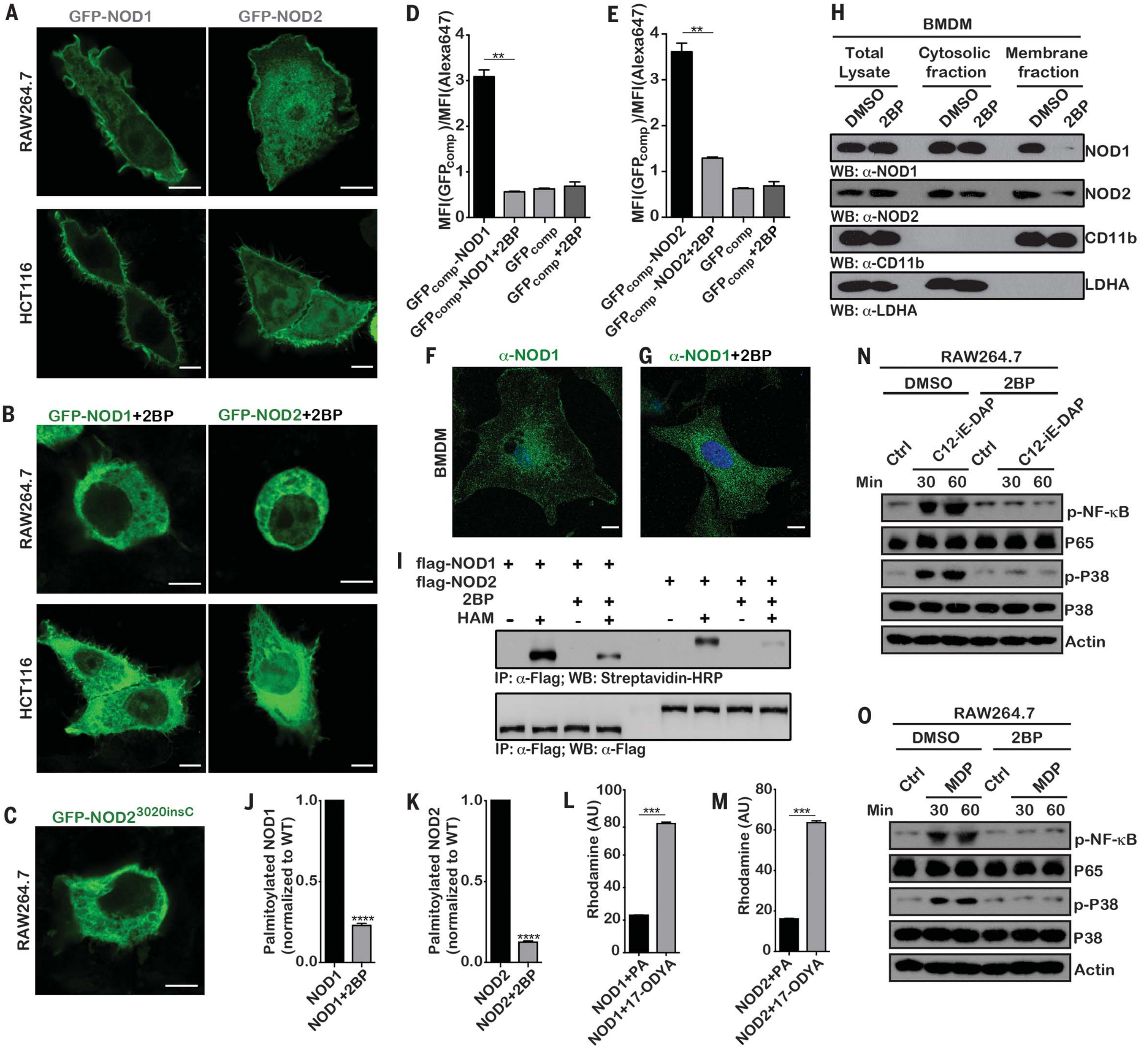


Fig. 1. NOD1/2 S-palmitoylation is required for membrane association and agonist-triggered signaling. (**A** and **B**) Fluorescence microscopy of RAW264.7 and HCT116 cells expressing GFP-NOD1 or GFP-NOD2 in the presence of (A) dimethyl sulfoxide (DMSO) or (B) 50 μM 2BP. (**C**) Fluorescence microscopy of RAW264.7 cells expressing the GFP-NOD2 CD variant 3020insC. (**D** and **E**) Plasmalemmal GFP-fragment complementation assay (GFP$_{comp}$ fluorescence) for (D) NOD1 and (E) NOD2. (**F** and **G**) Endogenous NOD1 [green; 4′,6-diamidino-2-phenylindole (DAPI) in blue] in mouse BMDM treated with (F) DMSO or (G) 2BP. (**H**) Total, cytosolic, and membrane fractions of BMDM pretreated with DMSO or 2BP (100 μM) were immunoblotted with antibodies directed against NOD1, NOD2, a cytosol marker (LDHA), or a PM (CD11b) marker. (**I**) S-palmitoylation levels of flag-NOD1/2 expressed in HEK293T cells in the presence of HAM or 2BP (100 μM). (**J** and **K**) Quantification of the 2BP treatment effect on (J) NOD1 and (K) NOD2 S-palmitoylation of immunoblots as in (I). (**L** and **M**) Covalent attachment of 17-ODYA to (L) flag-NOD1 or (M) flag-NOD2 expressed in HEK293T cells. (**N** and **O**) p65 and p38 phosphorylation levels in response to C12-iE-DAP- (250 ng ml^{-1}) or MDP (500 ng ml^{-1}) in RAW264.7 cells. Data in (H), (I), (N), and (O) are representative of three independent experiments, and data in (D), (E), and (J) to (M) represent the mean ± SEM of triplicate samples. ns, not significant; $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$, Student's *t* test; AU, arbitrary units. Scale bars, 5 μm.

that NOD1/2 are palmitoylated and that this posttranslational modification is required for optimal sensing of bacterial peptidoglycans and activation of NF-κB signaling.

Multiple cysteine residues in NOD1/2 are S-acylated

On the basis of evolutionary conservation, the crystal structure of the *Oryctolagus cuniculus* NOD2 protein (PDB:5IRL), and the loss of function in the NOD2^{C1033} truncation mutant protein, we created a series of NOD1 and NOD2 Cys to Ser mutants (NOD1^{C952S}, NOD1^{C557S}, NOD1^{C567S}, NOD1^{C558,567,952S}; NOD2^{C395S}, NOD2^{C1033S}, and NOD2^{C395,1033S}) (fig. S2, A and B), and their S-palmitoylation levels were characterized by using confocal microscopy and ABE assays. Similar to that observed for the CD variant NOD2^{3020insC} (Fig. 1C and fig. S1C), all of these single- or multiple-cysteine GFP-NOD1/2 mutants displayed cytosolic localization in HEK293 cells and RAW264.7 macrophages (Fig. 2A and fig. S2, C to E) and decreased palmitoylation levels (Fig. 2, B and C, and fig. S2, F and G). Palmitoylation of the NOD1 triple mutant NOD1^{C558,567,952S} (Fig. 2, B and C) and the NOD2 double-Cys mutant NOD2^{C395,1033S} (fig. S2, F and G) was further reduced by more than 50% as compared with the single NOD1/2 Cys to Ser mutants. The lack of S-palmitoylation did not affect the stability of the mutant proteins (fig. S2H). Similar to that observed in 2BP-treated RAW264.7 and HCT116 cells expressing wild-type (WT) NOD1/2 (Fig. 1H and fig. S1, M to O), the NOD1^{C558,567,952S} and NOD2^{C395,1033S} variant proteins were depleted from the membrane fraction in untreated HCT116 cells (Fig. 2D and fig. S2I) and displayed reduced reconstituted GFP fluorescence (Fig. 2E and fig. S2J) and impaired agonist-induced NF-κB–dependent luciferase activity (Fig. 2F and fig. S2, K to M).

To demonstrate that membrane recruitment of NOD1/2 is required for activation of the NF-κB pathway, we made use of an S-palmitoylation–dependent PM targeting sequence of the neuronal growth cone protein GAP43 (amino acids 1 to 11, MLCCMRRTKQV;

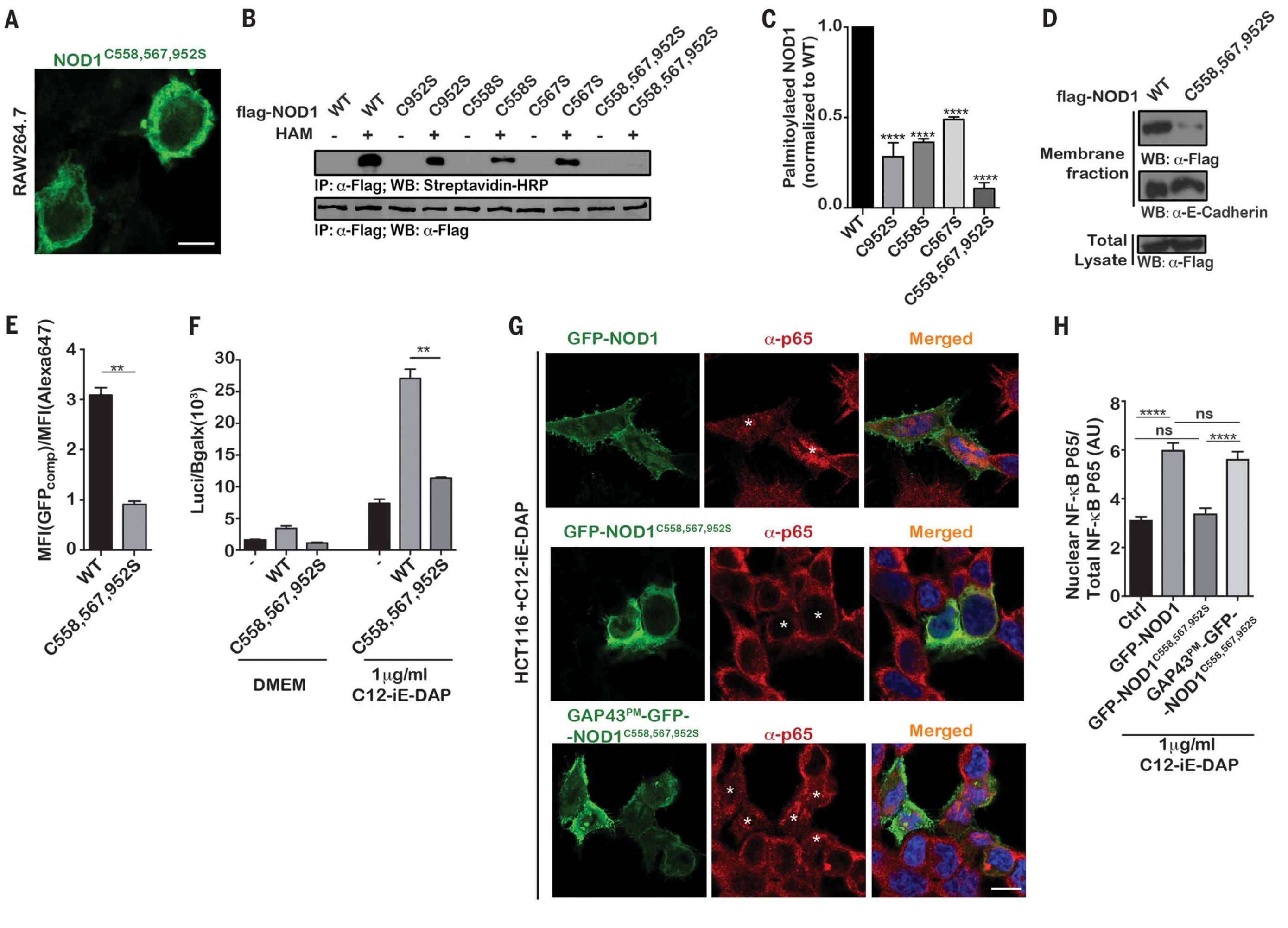


Fig. 2. Multiple cysteine residues in NOD1 can be acylated. (**A**) Fluorescence microscopy of RAW264.7 cells expressing a GFP-tagged S-palmitoylation–deficient NOD1 mutant (NOD1^{C558,567,952S}). (**B**) Streptavidin blot detection of NOD1 S-palmitoylation–deficient mutants. (**C**) Quantification of NOD1 S-palmitoylation–deficient mutants in (B). (**D**) Membrane fractions (or total lysate) of HCT116 cells expressing flag-WT or S-palmitoylation–deficient NOD1^{C558,567,952S} were immunoblotted for NOD1 or the PM marker E-cadherin. (**E**) Reconstituted plasmalemma-associated GFP_{comp} fluorescence for NOD1^{C558,567,952S}. (**F**) HEK293T cells were transfected with an NF-κB reporter β-Gal-SV40 expression plasmid and NOD1 overexpression plasmids. Luciferase assays were performed after stimulation with C12-iE-DAP (1 μg ml^{-1}). (**G**) NF-κB nuclear translocation in HCT116 cells transiently transfected with GFP-WT NOD1, S-palmitoylation–deficient GFP-NOD1^{C558,567,952S} or GAP43 (amino acids 1 to 11) tagged GFP-NOD1^{C558,567,952S} and stimulated with 500 ng ml^{-1} C12-iE-DAP. Cells were stained with antibody to NF-κB p65 and DAPI. White asterisks indicate the nuclei location in transfected cells. (**H**) Quantification of NF-κB nuclear translocation, relative to total NF-κB p65 ($n \geq 30$ cells for each condition). Data in (B) and (D) are representative of three independent experiments, and data in (C), (E), and (F) represent the mean ± SEM of triplicate samples. ns, not significant; $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, Student's *t* test. Scale bars, 5 μm.

hereafter denoted GAP43^{PS}) (*27*). Fluorescence microscopy of HCT116 cells expressing GFP-NOD1, GFP-NOD1^{C558,567,952S}, and GAP43^{PS}-GFP-NOD1^{C558,567,952S} revealed that the GAP43 S-palmitoylation motif could rescue the localization of the S-palmitoylation–deficient NOD1 mutant (Fig. 2G and fig. S2N, bottom) and induced the nuclear translocation of the NF-κB p65 (RelA) subunit (Fig. 2H and fig. S2N, bottom) and consequently induced NOD1 agonist NF-κB–dependent luciferase activity (fig. S2O). Thus, S-palmitoylation of NOD1/2 is required for their proper membrane targeting and the activation of immune signaling pathways.

ZDHHC5 is necessary for NOD1 and NOD2 S-palmitoylation

BirA-tagged proteins were generated (fig. S3A) and used in a BioID screen (*28*) to identify NOD1- and NOD2-interacting proteins (tables S1 and S2). As expected, a number of plasmalemmal and endomembrane proteins were identified as high-confidence NOD1- and NOD2-proximity interactors (fig. S3, B and C). Membrane proteins and membrane-associated polypeptides [as determined with Gene Ontology (GO) enrichment analysis: pantherdb.org GO:0016020] represented >85% (111 of 129) of the NOD1 and >80% (58 of 72) of the NOD2 high-confidence proximity interactors. Because NOD1/2 are membrane localized in the absence of agonists in 293 cells, the BioID screen was conducted in resting cells. Thus, known interactors of NOD1/2 such as tumor necrosis factor (TNF) receptor–associated factor 2 (TRAF2) or RIP2 were not found in our dataset. Among the NOD1/2 high-confidence proximity interactors was ZDHHC5 (fig. S3B) (*19*), a ubiquitously expressed integral membrane protein of the DHHC palmitoyltransferase (PAT) family. Similar to the NODs themselves, ZDHHC5 also localized to the PM and vesicular structures (fig. S3D). We thus hypothesized that ZDHHC5 could form a transient complex with NOD1/2. ZDHHC5 was detected in immunoprecipitates of both NOD1 and NOD2 (fig. S3, E and F), and NOD1/2 were detected in ZDHHC5 immunoprecipitates (fig. S3, G and H). ZDHHC5 did not interact with NLRP3, showing a degree of specificity for its interaction with NOD1/2 (fig. S3I). GFP-NOD1 and mCherry-ZDHHC5 displayed strong colocalization in RAW264.7 and HEK293 cells (fig. S3, J and K), and endogenous NOD1 colocalized with ZDHHC5 in *h*MDMs (fig. S3L). By contrast, GFP-NOD1 was largely cytosolic when coexpressed with a catalytically inactive ZDHHC5 mutant, ZDHHC5^{C134S} (*21*, *29*) (fig. S3M). Similar results were obtained for GFP-NOD2 (fig. S3, N and O).

Silencing the expression of *ZDHHC5* in RAW264.7 cells by means of small interfering RNA (siRNA) [validated with quantitative polymerase chain reaction (PCR)] (fig. S4A) or in

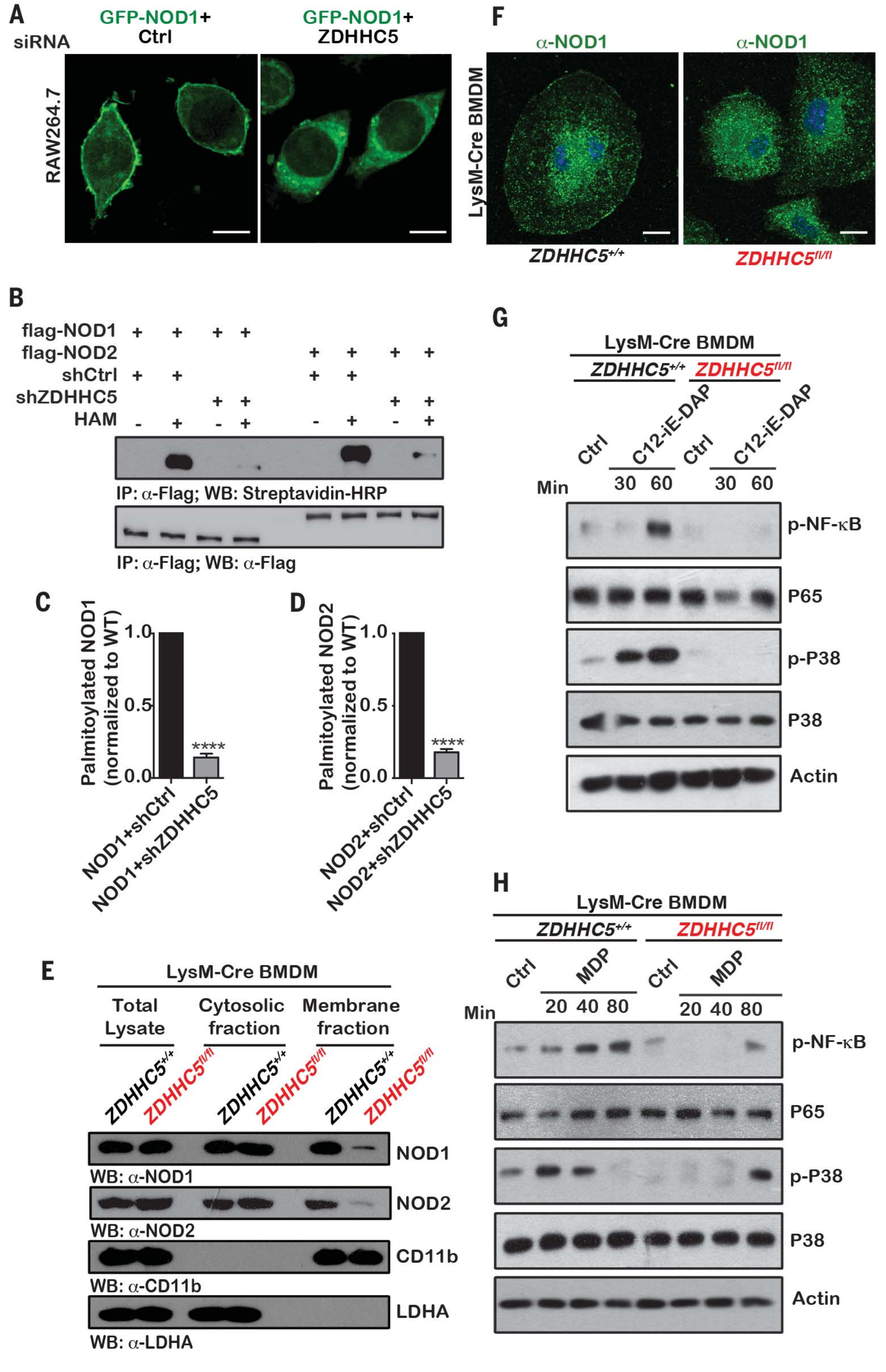


Fig. 3. The palmitoyl acyltransferase ZDHHC5 regulates NOD1 and NOD2 S-palmitoylation. (**A**) Representative images of GFP-NOD1 localization in RAW264.7 cells transfected with (left) nonspecific scramble RNA or (right) siRNAs targeting mouse *ZDHHC5*. (**B**) NOD1 and NOD2 S-palmitoylation levels in WT and *ZDHHC5* KO cells. (**C** and **D**) Densitometric analysis of immunoblots as in (B). (**E**) Total, cytosolic, and membrane fractions of BMDM cells from LysM-Cre/*ZDHHC5*$^{+/+}$ and LysM-Cre/*ZDHHC5*$^{fl/fl}$ mice were immunoblotted with antibodies directed against NOD1/2, a cytosol marker (LDHA), and the PM marker CD11b. (**F**) Confocal microscopic analysis of BMDM cells derived from LysM-Cre/*ZDHHC5*$^{+/+}$ or LysM-Cre/*ZDHHC5*$^{fl/fl}$ animals. Endogenous NOD1, green; DAPI, blue. (**G** and **H**) Effect of ZDHHC5 KO on C12-iE-DAP–induced (250 ng ml^{-1}) or MDP-induced (500 ng ml^{-1}) p65 and p38 kinase phosphorylation in mouse BMDM cells from LysM-Cre/*ZDHHC5*$^{+/+}$ and LysM-Cre/*ZDHHC5*$^{fl/fl}$ mice. Data in (B), (E), (G), and (H) are representative of three independent experiments, and data in (C) and (D) represent the mean ± SEM of triplicate samples. ****P < 0.0001, Student's t test. Scale bars, 10 μm.

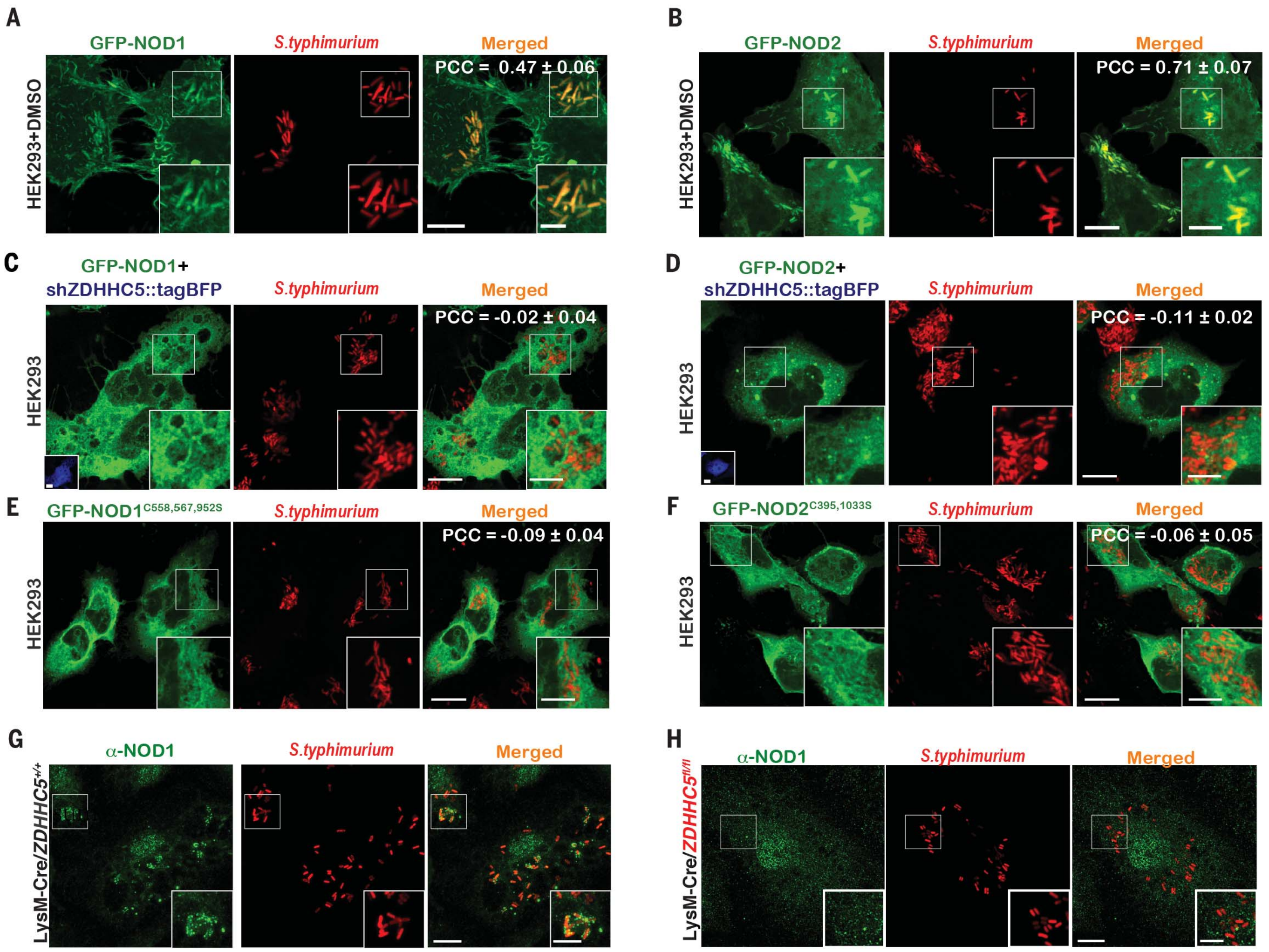


Fig. 4. ZDHHC5-mediated S-palmitoylation is indispensable for NOD1/2 recruitment to bacteria containing phagosomes. (**A** to **F**) Confocal microscopy of NOD1 and NOD2 in *Salmonella*-infected HEK293 cells. Cells transiently transfected with [(A) and (C)] GFP-NOD1, [(B) and (D)] GFP-NOD2, or [(C) and (D)] sh*ZDHHC5*::tagBFP (shown in the blue inset) and [(A) to (D)] infected with red fluorescent protein (RFP)–expressing *S. typhimurium*. [(E) and (F)] Confocal microscopy of *Salmonella*-infected HEK293 cells transiently expressing the indicated (E) S-palmitoylation deficient GFP-NOD1or (F) GFP-NOD2 mutants. Colocalization of NOD1/2 and *S. typhimurium* was described by the Pearson's correlation coefficient (PCC). (**G** and **H**) Confocal microscopy of *Salmonella*-infected ZDHHC5-deficient (LysM-Cre/*ZDHHC5*$^{fl/fl}$) or WT (LysM-Cre/*ZDHHC5*$^{+/+}$) BMDMs immunostained with α-NOD1 antibody (green). Scale bars, 10 μm and (insets) 5 μm.

HEK293 cells with a short hairpin RNA (shRNA) (fig. S4B) rendered NOD1/2 largely cytosolic (Fig. 3A and fig. S4, C and D), resembling the distribution observed after 2BP treatment (Fig. 1B and fig. S1B). Expression of the shRNA targeting *ZDHHC5* also caused profound defects in NOD1 and NOD2 palmitoylation (Fig. 3, B to D), as assessed with the ABE assay. By contrast, the related PAT ZDHHC20 did not associate with NOD1/2 (fig. S5, A and B), and silencing of its expression did not affect NOD1/2 S-palmitoylation (fig. S5, D to G). Silencing ZDHHC5 expression in RAW264.7 cells also abrogated C12-iE-DAP–induced NF-κB activation as assessed with phospo-p65 and phospho-p38 immunoblotting (fig. S4E), similar to that observed in HEK293T cells (fig. S4, F and G).

A myeloid cell lineage–specific conditional *ZDHHC5* knockout (KO) animal model was generated by crossing mice carrying a floxed-*ZDHHC5* allele (*ZDHHC5*$^{fl/fl}$) with myelomonocytic cell–specific Lysozyme M (LysM)–Cre Recombinase mice (fig. S4, H and I). Backcrossing of LysM-Cre/ZDHHC5$^{fl/+}$ mice led to homozygous ablation of ZDHHC5 in the mouse myeloid cell lineage (LysM-Cre/*ZDHHC5*$^{fl/fl}$). Genotyping by means of PCR confirmed the presence of LoxP and LysM-Cre in LysM-Cre/*ZDHHC5*$^{fl/fl}$ mice (fig. S4J). Immunoblot analysis and immunofluorescence microscopy confirmed the depletion of ZDHHC5 protein expression in LysM-Cre/*ZDHHC5*$^{fl/fl}$ BMDM (fig. S4, K and L). These mice exhibited no obvious phenotype under standard laboratory conditions. Cellular fractionation of WT and *ZDHHC5* KO *m*BMDMs confirmed that NOD1/2 membrane association was substantially reduced in the absence of ZDHHC5 (Fig. 3E and fig. S4M). In contrast to that observed in WT mice, endogenous NOD1 was predominantly cytosolic in ZDHHC5 KO *m*BMDMs (Fig. 3F), and ZDHHC5 KO *m*BMDMs displayed an impaired NOD1/2–dependent activation of NF-κB and p38 MAPK signaling in response to C12-iE-DAP (Fig. 3G) or MDP (Fig. 3H), despite similar NOD1/2 protein expression levels (Fig. 3E).

NOD1/2 S-palmitoylation levels increased by more than twofold upon stimulation with C12-iE-DAP or MDP (fig. S6, A to C) and correlated with additional recruitment of NOD1/2 proteins to peripheral membranes (fig. S6, D to J). Thus, subcellular fractionation of

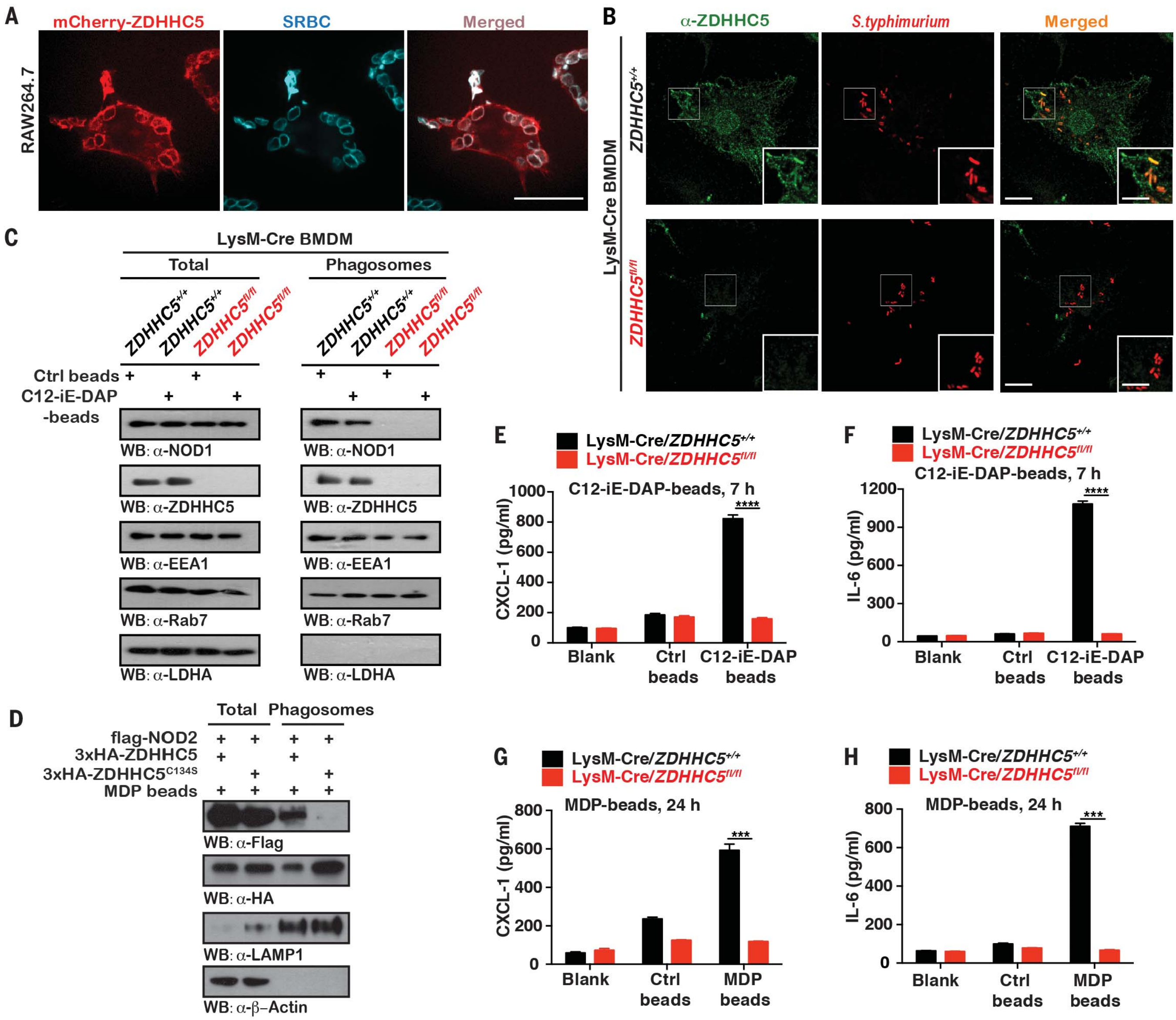


Fig. 5. ZDHHC5 is recruited to bacteria-containing phagosomes. (**A**) Confocal microscopy of RAW264.7 cells expressing mCherry-ZDHHC5 challenged with red blood cells labeled with Cy5-conjugated IgG (blue). (**B**) Confocal microscopy of wild-type (LysM-Cre/*ZDHHC5*$^{+/+}$) and ZDHHC5-deficient (LysM-Cre/*ZDHHC5*$^{fl/fl}$) BMDMs immunostained with α-ZDHHC5 antibody (green) and infected with RFP–*S. typhimurium*, as in Fig. 4G. (**C**) Western blot analysis of C12-iE-DAP–coated magnetic beads containing phagosomes isolated from ZDHHC5-deficient (LysM-Cre/*ZDHHC5*$^{fl/fl}$) or WT (LysM-Cre/*ZDHHC5*$^{+/+}$) BMDMs, respectively. (**D**) Western blot analysis of MDP-coated magnetic beads containing phagosomes isolated from HEK293T cells coexpressing flag-NOD2 with 3xHA-ZDHHC5 or catalytically inactive 3xHA-ZDHHC5^{C134S}. (**E** to **H**) ZDHHC5-deficient (LysM-Cre/*ZDHHC5*$^{fl/fl}$) or wild-type (LysM-Cre/*ZDHHC5*$^{+/+}$) BMDMs were incubated with C12-iE-DAP– or MDP-coated beads. CXCL-1 and IL-6 in medium were quantified by means of enzyme-linked immunosorbent assay. Data in (C) and (D) are representative of three independent experiments, and data in (E) to (H) represent the mean ± SEM of triplicate samples. ns, not significant; $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, Student's *t* test; AU, arbitrary units. Scale bars, 10 μm.

HCT116 cells and primary *m*BMDMs stimulated under similar conditions yielded similar results (fig. S6, D to J). This agonist-mediated increase in S-palmitoylation was dependent on ZDHHC5 because cells failed to respond to NOD agonists in its absence (fig. S6, K to M). Thus, ZDHHC5 mediates NOD1/2 S-palmitoylation, and this modification is required for membrane association and downstream signaling in a physiologically relevant context.

ZDHHC5-mediated palmitoylation regulates NOD1/2 function

GFP-NOD1 and GFP-NOD2 are localized to *Salmonella*-containing vacuoles (SCV) in *Salmonella typhimurium*–infected cells (Fig. 4, A and B, and fig. S7, A and B) (*6*), and SCV localization was markedly reduced with 2BP treatment (fig. S7, C and D) or ZDHHC5 knockdown (Fig. 4, C and D). Consistent with this observation, palmitoylation-deficient NOD1/2 mutants failed to localize to SCVs (Fig. 4, E and F, and fig. S7, E and F), and levels of flag-NOD2 protein associated with phagosomes containing MDP-coated beads were substantially decreased by 2BP treatment (fig. S7G).

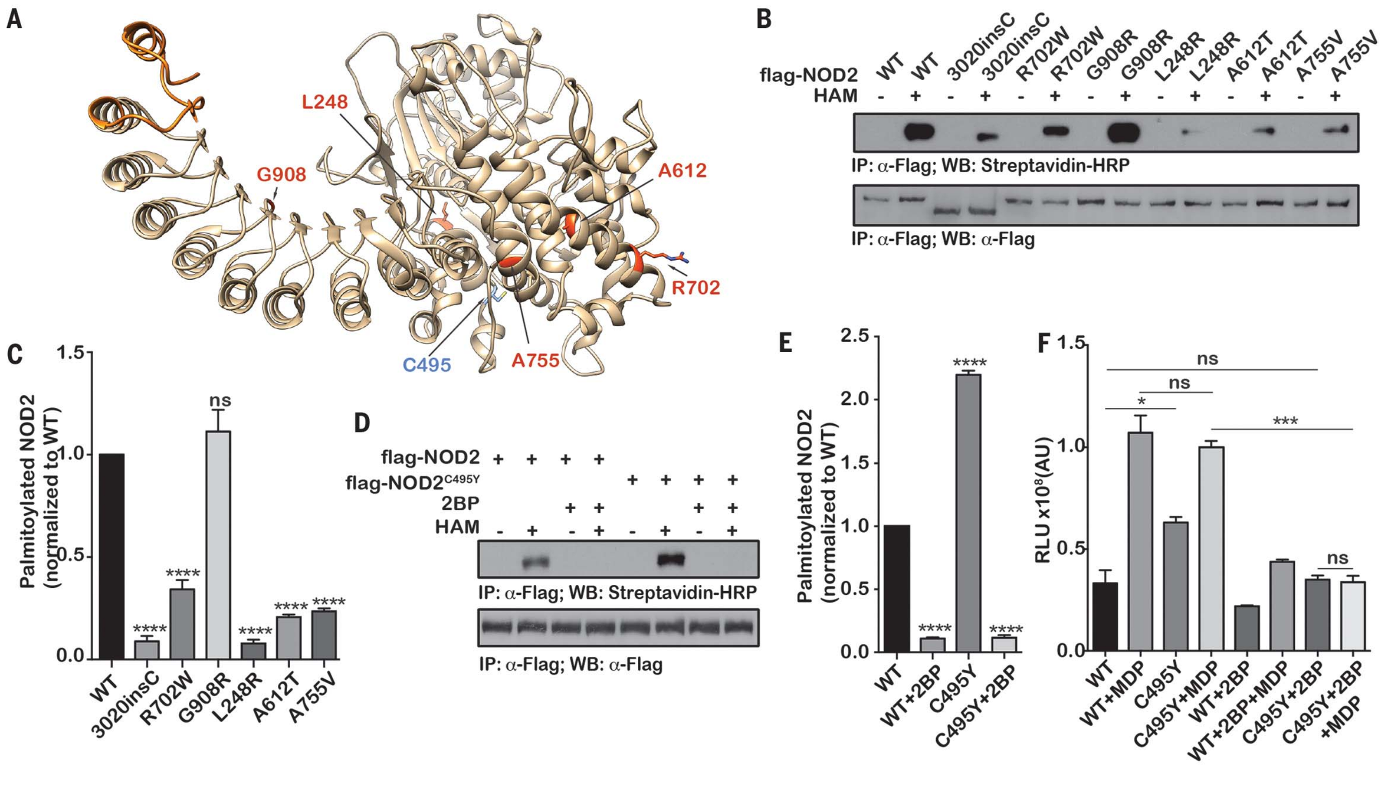


Fig. 6. NOD2 coding variants manifest aberrant S-palmitoylation. (**A**) Ribbon representation of CD (3020insC, orange; R702, G908, L248, A612, and A755, red) and EOS (C495Y, blue) mutations on a *h*NOD2 3D structure model. (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example, R702W indicates that arginine at position 702 was replaced by tryptophan.) (**B**) S-palmitoylation levels of the indicated NOD2 variants (3020insC, R702, G908, L248, A612, and A755) expressed in HEK293T cells with or without HAM. (**C**) Quantification of S-palmitoylation of NOD2 variants of immunoblots as in (B). (**D**) Effect of 2BP (50 μM) on S-palmitoylation of the NOD2 EOS variant NOD2^{C495Y}. (**E**) Quantification of S-palmitoylation level of NOD2^{C495Y} in (D). (**F**) NF-κB–dependent luciferase reporter gene activity in HEK293 cells expressing GFP-NOD2^{C495Y} treated with MDP (2 μg ml^{-1}) and DMSO or 2BP (50 μM). Luciferase activity normalized as in Fig. 1G. Data in (C), (E) and (F) are the mean ± SEM of at least triplicate samples. ns, not significant; *$P < 0.05$, ***$P < 0.001$, ****$P < 0.0001$, Student's *t* test;. RLU, relative luminescence units; AU, arbitrary units.

Consistent with these observations, endogenous NOD1 was associated with phagosomes in WT *m*BMDMs (Fig. 4G) but not KO *m*BMDMs (Fig. 4H).

Similar to NOD1/2 (Fig. 4, A and B), overexpressed mCherry-ZDHHC5 (Fig. 5A), GFP-ZDHHC5 (fig. S8A), and endogenous ZDHHC5 were recruited to large phagosomes (fig. S8B) and SCVs in infected WT *m*BMDMSs and *h*MDMs (Fig. 5B and fig. S8C) (*1*, *6*). These observations were in line with the NODs and ZDHHC5 co-fractionating with early and late endosomal markers. Consistently, endogenous ZDHHC5 and NOD1 were found on C12-iE-DAP–coated bead–containing phagosomes isolated from *m*BMDMs after asynchronous phagocytosis (Fig. 5C). By contrast, C12-iE-DAP bead phagosomes isolated from LysM-Cre/*ZDHHC5$^{fl/fl}$* *m*BMDMs were devoid of NOD1 but were positive for EEA1 and Rab7 (Fig. 5C). Furthermore, overexpressed ZDHHC5 was associated with LAMP1-positive phagosomes generated by the internalization of C12-iE-DAP– or MDP-coated beads (Fig. 5D and fig. S8D). The recruitment of ZDHHC5 to phagosomes did not require its transferase activity (Fig. 5D). However, overexpression of catalytically inactive ZDHHC5 prevented the recruitment of NOD2 to MDP-coated bead–containing phagosomes (Fig. 5D), similar to cells treated with 2BP (fig. S7G).

Consistent with previous findings (*6*), when WT macrophages were incubated with C12-iE-DAP (Fig. 5, E and F) or MDP beads (Fig. 5, G and H), CXCL-1 (Fig. 5, E and G), and interleukin-6 (IL-6) (Fig. 5, F and H) were produced. However, CXCL-1 and IL-6 production by ZDHHC5-deficient *m*BMDMs was markedly lower than that in littermate control cells, despite internalizing more beads (fig. S8E). Together, these data indicate that ZDHHC5 is constitutively localized to phagosomes and that its catalytic activity is required for NOD1/2 phagosomal recruitment and NOD1/2–dependent signaling.

NOD2 S-palmitoylation and disease relevance

NOD1 and *NOD2* polymorphisms are associated with CD, ulcerative colitis (UC), Blau syndrome, Behcet's syndrome, early-onset sarcoidosis (EOS), and atopic diseases (*1*, *2*). CD-associated NOD2 loss-of-function mutations result in reduced NF-κB activity upon stimulation with MDP (*6*), whereas EOS and Blau syndrome are associated with apparent NOD2 gain-of-function mutations (*30*, *31*). The three most common nonsynonymous NOD2 variants (R702W, G908R, and 3020insC) (Fig. 6A and fig. S9A) account for roughly 80% of all NOD2-associated variants found in CD cases (*1*) and retain the ability to bind to MDP with nanomolar affinity (*10*).

We found that a number of CD-associated NOD2 mutant proteins (3020insC, R702W, L248R, A612T, A755V, and R1019stop) displayed a ≈70 to 90% reduction in S-palmitoylation levels (Fig. 6, B and C, and fig. S9, B and C). As expected, decreased palmitoylation strongly correlates with a more pronounced cytosolic distribution (Fig. 1C and fig. S9D). We thus surmise that a reduction in the S-palmitoylation of these NOD2 variants leads to membrane dissociation and hence an inability to respond to MDP, despite their nanomolar affinity for this compound (*10*). By contrast, the EOS-associated

NOD2 gain-of-function mutant C495Y displayed a twofold increase in S-palmitoylation compared with that of WT NOD2 (Fig. 6, D and E). NOD2 C495Y palmitoylation and associated NF-κB hyperactivity could be counterbalanced by the use of 2BP (Fig. 6F). These data indicate a strong correlation between defects in NOD2 (and potentially NOD1) palmitoylation and several different human pathological conditions.

Discussion

We have established that S-palmitoylation of both NOD1 and NOD2 is required for optimal membrane targeting and hence for proper signaling upon cognate peptidoglycan detection. Membrane association of NOD1 and NOD2 is vital for their function (*1*, *7*, *8*, *16*), as clearly illustrated by the CD-associated NOD2 loss-of-function variant 3020insC, which fails to associate with membranes and, as a consequence, is unable to detect bacterial intrusion (*6*). We provide evidence that NOD1 and NOD2 S-palmitoylation is required not only for steady-state membrane association but also for ligand-induced signaling. Accordingly, S-palmitoylation–deficient mutants of NOD1 and NOD2 are mislocalized and lose their ability to induce NF-κB signaling in response to C12-iE-DAP or MDP. Furthermore, we found that ZDHHC5 is required for S-palmitoylation of NOD1 and NOD2, and both the presence of ZDHHC5 at the site of bacterial entry and its enzymatic activity are necessary for proper recruitment of NOD1/2 to the bacterial entry site and phagosomes. How ZDHHC5 is recruited to the site of bacterial entry remains an open question.

We conclude that a palmitoylation-dependent local accumulation of NODs, and their subsequent exposure to bacterial by-products, ensures a compartmentalized and efficient signaling response. Consequently, cells can respond to very low concentrations of bacterial products. We propose that NOD1 and NOD2, together with ZDHHC5 and various transporters (such as SLC15A3), form specialized platforms for pathogen sensing. We anticipate that our findings will advance the understanding and ultimately the treatment of NOD-driven inflammatory diseases, including NOD-dependent autoimmune diseases and chronic bacterial infections.

REFERENCES AND NOTES

1. D. J. Philpott, M. T. Sorbara, S. J. Robertson, K. Croitoru, S. E. Girardin, *Nat. Rev. Immunol.* **14**, 9–23 (2014).
2. R. Caruso, N. Warner, N. Inohara, G. Núñez, *Immunity* **41**, 898–908 (2014).
3. J. P. Hugot *et al*., *Nature* **411**, 599–603 (2001).
4. Y. Ogura *et al*., *Nature* **411**, 603–606 (2001).
5. T. A. Kufer, E. Kremmer, A. C. Adam, D. J. Philpott, P. J. Sansonetti, *Cell. Microbiol.* **10**, 477–486 (2008).
6. N. Nakamura *et al*., *Nature* **509**, 240–244 (2014).
7. N. Barnich, J. E. Aguirre, H. C. Reinecker, R. Xavier, D. K. Podolsky, *J. Cell Biol.* **170**, 21–26 (2005).
8. A. T. Irving *et al*., *Cell Host Microbe* **15**, 623–635 (2014).
9. J. H. Park *et al*., *J. Immunol.* **178**, 2380–2386 (2007).
10. A. K. Schaefer *et al*., *Biochemistry* **56**, 4445–4448 (2017).
11. S. W. Brubaker, K. S. Bonham, I. Zanoni, J. C. Kagan, *Annu. Rev. Immunol.* **33**, 257–290 (2015).
12. A. Fukazawa *et al*., *PLOS Pathog.* **4**, e1000228 (2008).
13. T. A. Kufer, E. Kremmer, D. J. Banks, D. J. Philpott, *Infect. Immun.* **74**, 3115–3124 (2006).
14. C. Stevens *et al*., *Gut* **62**, 695–707 (2013).
15. S. Lipinski *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21426–21431 (2012).
16. A. M. Keestra *et al*., *Nature* **496**, 233–237 (2013).
17. K. S. Bonham, J. C. Kagan, *Cell Host Microbe* **15**, 523–525 (2014).
18. M. E. Linder, R. J. Deschenes, *Nat. Rev. Mol. Cell Biol.* **8**, 74–84 (2007).
19. J. Greaves, L. H. Chamberlain, *Trends Biochem. Sci.* **36**, 245–253 (2011).
20. H. Jiang *et al*., *Chem. Rev.* **118**, 919–988 (2018).
21. M. S. Rana *et al*., *Science* **359**, eaao6326 (2018).
22. J. S. Yount, M. M. Zhang, H. C. Hang, *Curr. Opin. Chem. Biol.* **17**, 27–33 (2013).
23. M. D. Resh, *Methods* **40**, 191–197 (2006).
24. D. Kamiyama *et al*., *Nat. Commun.* **7**, 11046 (2016).
25. R. C. Drisdel, J. K. Alexander, A. Sayeed, W. N. Green, *Methods* **40**, 127–134 (2006).
26. B. R. Martin, B. F. Cravatt, *Nat. Methods* **6**, 135–138 (2009).
27. Y. Fukata *et al*., *J. Cell Biol.* **202**, 145–161 (2013).
28. K. J. Roux, D. I. Kim, M. Raida, B. Burke, *J. Cell Biol.* **196**, 801–810 (2012).
29. Y. Li, B. R. Martin, B. F. Cravatt, S. L. Hofmann, *J. Biol. Chem.* **287**, 523–530 (2012).
30. B. Zurek, M. Proell, R. N. Wagner, R. Schwarzenbacher, T. A. Kufer, *Innate Immun.* **18**, 100–111 (2012).
31. R. Parkhouse, J. P. Boyle, T. P. Monie, *FEBS Lett.* **588**, 3382–3389 (2014).

ACKNOWLEDGMENTS

The authors thank S. Grinstein for advice with experimental design, valuable discussions, and critically reading the manuscript as well as P. Bilan for assistance with development of experimental methods and tools. We are also thankful for the technical support by the Core Facilities, Zhejiang University School of Medicine. **Funding:** This study was supported by grants from the Natural Science Foundation of China (NSFC) (31770938 and 91854113) and the Key Program of Zhejiang Provincial Natural Science Foundation of China (LZ16C050001) to D.N.; Q.S. is supported by grants from NSFC (31771525 and 91754113) and Ministry of Science and Technology of the People's Republic of China (2017YFA0503402). G.D.F is supported by Canadian Institutes of Health Research (CIHR) (MOP-133656 and PJT-166010) and the J. P. Bickell Foundation. J.H.B. is supported by operating grants from the CIHR (FDN#154329) and holds the Pitblado Chair in Cell Biology. A.K. was supported by a Foundation Grant from CIHR (FND#143203). A.M.M. is funded by the CIHR and the Leona M. and Harry B. Helmsley Charitable Trust. **Author contributions:** D.N., Q.S., B.R., E.C., and G.D.F. designed the research; Y.Lu, Y.Z., E.C., C.Z., Y.Y., X.W., J.S.C, Y.M., and X.C. performed the experiments; D.N., Q.S., B.R., and G.D.F. analyzed the results; D.N., G.D.F., and Q.S. wrote the paper; Y.Li. generated *h*DMCs. B.R., J.H.B., S.E.G., G.D.F., A.K., N.W., and H.H. edited and commented on the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Proteomic data have been deposited at massive.ucsd.edu (ID MSV000082079). All DNA constructs and the ZDHHC5 KO mouse model described in this study are available upon request. All other data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/460/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 and S2
References (*32*–*47*)

View/request a protocol for this paper from *Bio-protocol*.

2 July 2018; resubmitted 19 July 2019
Accepted 3 October 2019
10.1126/science.aau6391

STRUCTURAL BIOLOGY

Structural basis for the docking of mTORC1 on the lysosomal surface

Kacper B. Rogala[1,2,3,4,5], Xin Gu[1,2,3,4,5], Jibril F. Kedir[1,2,3,4,5], Monther Abu-Remaileh[1,2,3,4,5], Laura F. Bianchi[1], Alexia M. S. Bottino[1], Rikke Dueholm[1], Anna Niehaus[1], Daan Overwijn[1], Ange-Célia Priso Fils[1], Sherry X. Zhou[1], Daniel Leary[4], Nouf N. Laqtom[1], Edward J. Brignole[5,6], David M. Sabatini[1,2,3,4,5]*

The mTORC1 (mechanistic target of rapamycin complex 1) protein kinase regulates growth in response to nutrients and growth factors. Nutrients promote its translocation to the lysosomal surface, where its Raptor subunit interacts with the Rag guanosine triphosphatase (GTPase)–Ragulator complex. Nutrients switch the heterodimeric Rag GTPases among four different nucleotide-binding states, only one of which (RagA/B•GTP–RagC/D•GDP) permits mTORC1 association. We used cryo–electron microscopy to determine the structure of the supercomplex of Raptor with Rag-Ragulator at a resolution of 3.2 angstroms. Our findings indicate that the Raptor α-solenoid directly detects the nucleotide state of RagA while the Raptor "claw" threads between the GTPase domains to detect that of RagC. Mutations that disrupted Rag-Raptor binding inhibited mTORC1 lysosomal localization and signaling. By comparison with a structure of mTORC1 bound to its activator Rheb, we developed a model of active mTORC1 docked on the lysosome.

The mTORC1 protein kinase controls growth in response to diverse environmental cues, such as nutrients and growth factors. Deregulated mTORC1 signaling is associated with many diseases, including some cancers and neurological disorders (*1*–*5*). Amino acids promote the translocation of mTORC1 to the surface of the lysosome, where it can interact with and be activated by the Rheb GTPase (*6*–*10*). mTORC1, composed of the core mTOR, Raptor, and mLST8 subunits, docks on the lysosome through the direct interaction of Raptor with the lysosome-associated Rag GTPase–Ragulator complex (*11*, *12*). Through their C-terminal roadblock domains (CRDs), the Rag GTPases form heterodimers consisting of RagA or RagB bound to RagC or RagD (*13*, *14*). The obligate heterodimeric nature of the Rags allows crosstalk between their GTPase domains, which is necessary for mTORC1 signaling to respond quickly to changes in nutrient levels (*15*). Ragulator consists of five subunits and is necessary for targeting the Rag GTPases to the lysosomal surface (*11*).

As in other GTPases, the GTPase domains of the Rags consist of a network of secondary structural elements, known as switches, that undergo conformational changes upon the exchange or hydrolysis of bound guanosine di- or triphosphate (GDP or GTP), respectively (*16*). Under the control of several nutrient-regulated GTPase-activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs), such as GATOR1 (*17*, *18*), FLCN-FNIP (*19*, *20*), and SLC38A9 (*21*), the Rag heterodimer can oscillate among four nucleotide configurations, only one of which (RagA/B•GTP–RagC/D•GDP) represents nutrient sufficiency and interacts with mTORC1 (*8*, *19*). To understand how mTORC1 discriminates between the Rag nucleotide states, we used cryo–electron microscopy (cryo-EM) to determine the structure of the Raptor subunit of mTORC1 bound to the Rag-Ragulator complex. This structure not only sheds light on the conformations of the Rag GTPases that underlie nutrient sensing by mTORC1, but also allows us to develop a structural model of mTORC1 docked on the lysosome.

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA. [2]Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Koch Institute for Integrative Cancer Research, Cambridge, MA 02139, USA. [4]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. [5]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [6]MIT.nano, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Corresponding author. Email: sabatini@wi.mit.edu

Reconstitution of the Raptor-Rag-Ragulator supercomplex

Because the interactions that promote the association of mTORC1 with the lysosomal surface underlie signaling events and are thus transient, we built the Raptor-Rag-Ragulator complex using a bottom-up approach instead of trying to isolate an intact complex from human cells. We generated the RagA-RagC heterodimer and the pentameric Ragulator complex in bacteria, and produced Raptor in a human embryonic kidney (HEK)–293 cell expression system (see supplementary materials). To stabilize the interaction of Raptor with the Rag GTPases, we introduced two point mutations into RagC (S75N, T90N) that are found in patients with follicular lymphoma (*22*). Each of these mutations independently stabilizes the GDP-bound state of RagC and promotes the interaction of the heterodimer with Raptor (*22*), and we found that in combination they have additive effects. We used the wild-type version of RagA because we assumed that its slow GTPase rate (*15*) would keep it bound to GTP. Indeed, analysis of the RagA-RagC (S75N, T90N) heterodimer confirmed that it contained close to stoichiometric amounts of GTP and GDP (fig. S1A). By mixing together the Rag GTPase heterodimer, Ragulator, and Raptor, we produced a fairly homogeneous supercomplex that was suitable for structural studies (Fig. 1A).

Cryo-EM structure of the Raptor-Rag-Ragulator supercomplex

In initial applications of the supercomplex to holey carbon microscopy grids, the particles strongly adsorbed to carbon, but after extensive optimization of blotting and freezing conditions, we obtained grids with a gradient of particle distributions. The edges of holes attracted many overlapping particles, whereas the centers had well-distributed individual complexes. The uneven particle distribution proved to be a major obstacle to many particle-picking algorithms, and only the Topaz and crYOLO pickers (*23*, *24*) discriminated well enough between particles on carbon and in holes to provide a uniform range of orientations and yield good two-dimensional classes (Fig. 1B).

After a series of refinements and particle sorting, we selected 112,037 particles to reconstruct the cryo-EM structure of the Raptor-Rag-Ragulator complex at 3.18 Å resolution (Fig. 1C and movie S1). The structure reveals the binding mode between Raptor and the Rag-Ragulator complex. The central region of Raptor (amino acids 376 to 844), which constitutes an α-solenoid (Fig. 1D), engages the Rag GTPases, whereas its N-terminal RNC (Raptor N-terminal conserved) domain and C-terminal WD40 domain do not. Although Ragulator makes extensive contacts with the CRDs of the Rags (fig. S1B), it does not directly interact with Raptor. Overall, Raptor and Rag-Ragulator form two elongated halves of the structure, positioned off-parallel to one another. The local resolution of the core of the supercomplex, formed by Rag-Raptor(α-solenoid), extends to ~2.3 Å resolution; the periphery of the structure was resolved to ~4.2 Å (fig. S9). Variability analysis revealed that such nonuniform resolution stems from the dynamics of the two halves of the complex, which swing with respect to one another. The highest amplitudes are observed at the tip of Ragulator and in the RNC and WD40 domains of Raptor (movie S2). The interface between Rag GTPase and CRD domains serves as a hinge region for the swinging movement of the Rag(CRD)-Ragulator substructure, which explains why capturing the full-length Rag-Ragulator with x-ray crystallography was likely an impossible task (*25*, *26*).

As in the structure of the Rag GTPases bound to GATOR1 (*27*), the relative positions

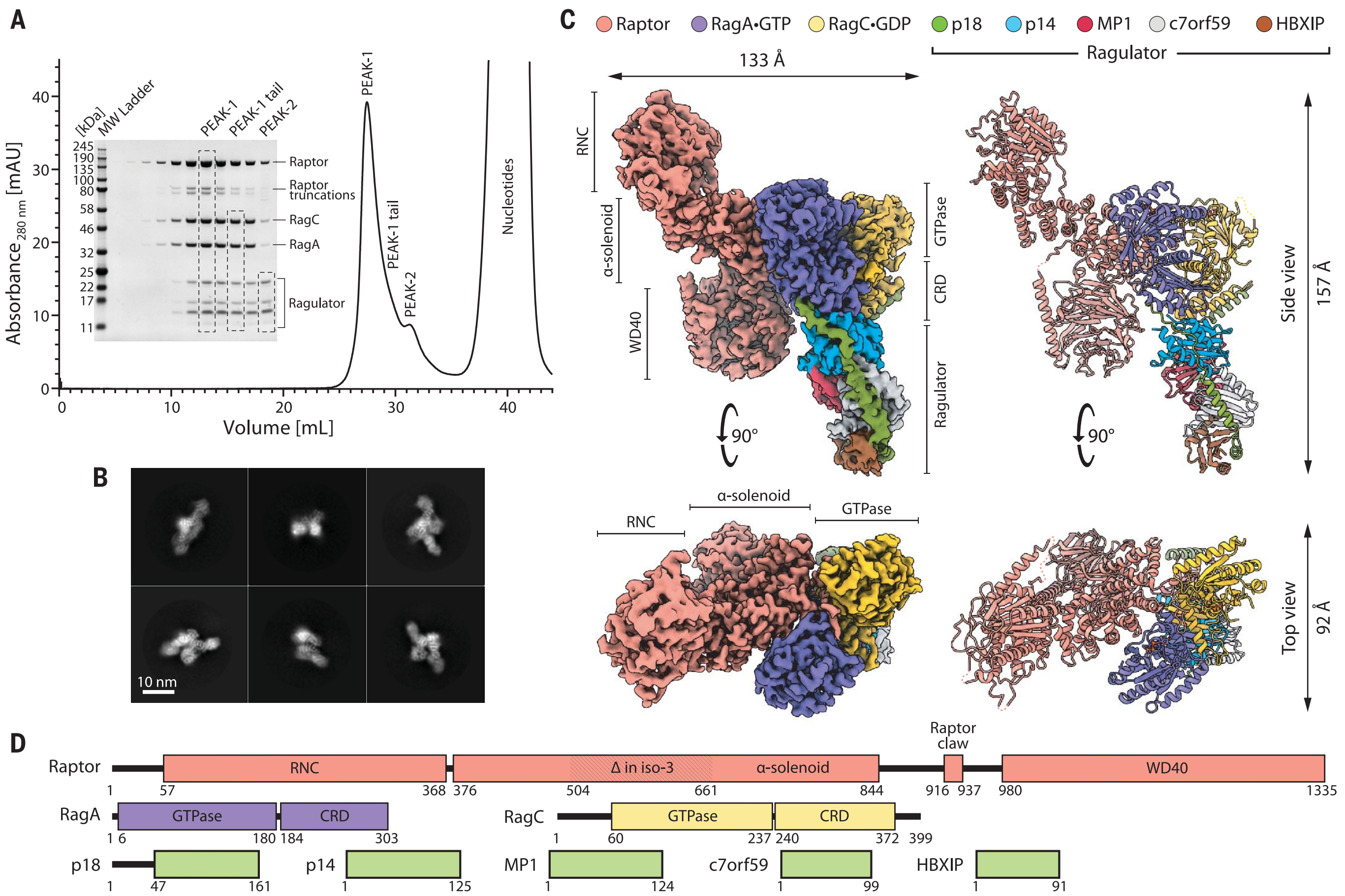


Fig. 1. Purification, assembly, and structure determination of the Raptor-Rag-Ragulator supercomplex. (**A**) Gel filtration profile and corresponding SDS–polyacrylamide gel electrophoresis of the reconstituted Raptor-Rag-Ragulator supercomplex as visualized with Coomassie Blue staining. The fully assembled complex (peak 1) partially overlaps with two subcomplexes: Rag-Ragulator (peak 1 tail) and Ragulator (peak 2). (**B**) Representative two-dimensional class averages of the Raptor-Rag-Ragulator supercomplex. (**C**) Cryo-EM structure of the supercomplex, determined to 3.2 Å resolution. Two orthogonal views of the experimental electron density (left) are shown next to corresponding views of the molecular model (right). (**D**) Domain organization of all components that make up the supercomplex.

of the GTPase domains in our current structure are altered with respect to those in previously reported structures of the Rag homologs in budding yeast, the Gtr proteins [PDB IDs 3R7W and 4ARZ (*28*, *29*)] (fig. S1C). In fact, the Gtr1•GTP-Gtr2•GDP active state would clash with Kog1 (the yeast homolog of Raptor) in a binding mode modeled on human Raptor-Rag. This suggests either that the structure of Kog1 adapts to accommodate the different Gtr structure or that the Gtr GTPases shift their conformation upon Kog1 binding. A structure of the yeast Gtr1-Gtr2-Kog1 complex will be needed to resolve the true situation among these possibilities.

The interaction of Raptor with RagA

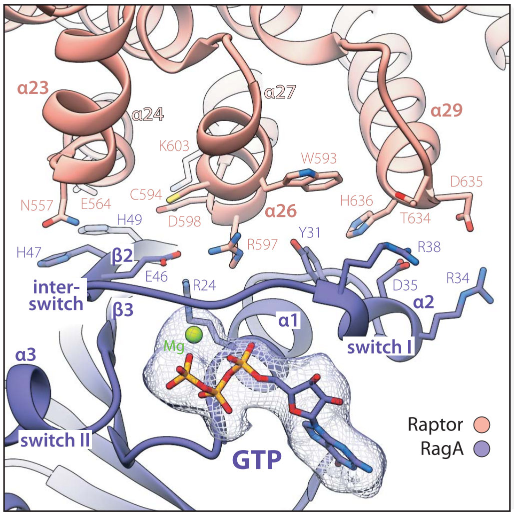
Raptor and RagA interact extensively with each other within the Raptor-Rag-Ragulator supercomplex. Three helices (α24, α26, α29) within a ~100–amino acid region (amino acids 546 to 650) of the Raptor α-solenoid (amino acids 376 to 844) directly engage the switch I–containing face of RagA and form a network of hydrogen bonds and salt bridges (Fig. 2A). The Glu^{564} and Asn^{557} residues of helix α24 of Raptor contact the RagA interswitch residues His^{49} and His^{47}. Arg^{597} of helix α26 forms extensive interactions by tightly fitting the pocket formed by Glu^{46}, Arg^{24}, and Tyr^{31} of RagA. These interactions are supported by hydrogen bonds between the backbone of His^{49} of RagA and Asp^{598} and Lys^{603} of Raptor. The third helix α29 of Raptor engages helix α2 of RagA, with a number of salt bridges between Thr^{634}, Asp^{635}, and His^{636} of Raptor and Arg^{38}, Arg^{34}, and Asp^{35} of RagA (Fig. 2A).

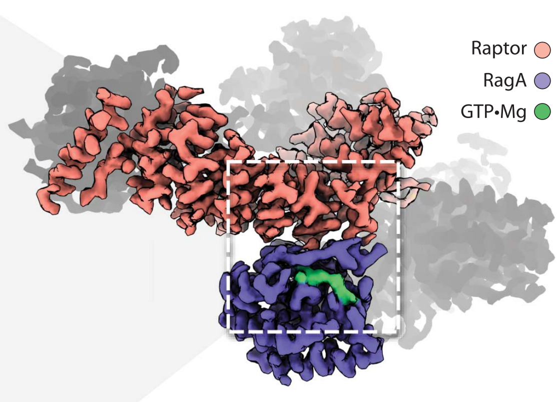
To test the importance of these contacts in mediating the Raptor-Rag interaction, we mutated interface residues on either Raptor or RagA. In Raptor, we disrupted either all three interacting helices involved in the interaction with one mutation per helix, or each of the three helices individually by introducing three mutations in each helix (Fig. 2B). All three helices, α24, α26, and α29, appear necessary for the binding of Raptor to the Rag heterodimer, with either mutations in all helices or mutations in helix α26 showing the strongest effect; this finding suggests that helix α26 plays a primary role in complex formation (Fig. 2C). None of the mutations affected the binding of Raptor to mTOR (fig. S2A).

Because the Raptor-RagA interface involves the critical switch I and interswitch regions of RagA, we only targeted, singly or in combination, residues on RagA that point away from the core of the protein (Fig. 2D) and validated that these mutations do not affect its capacity to bind GTP (fig. S2B). Consistent with the structural data, when expressed in RagA-RagB double-knockout (DKO) HEK-293T cells lacking RagA and its redundant homolog RagB (fig. S2C), none of the RagA mutants interacted with Raptor, whereas all still heterodimerized with RagC and bound to Ragulator, as assessed by coimmunoprecipitation of its p18/LAMTOR1 subunit (Fig. 2E).

On the basis of precedence with many small GTPases to which RagA is related (*29*, *30*), we predict that in the GDP-bound state, both switch I and strand β2 of the interswitch of

A





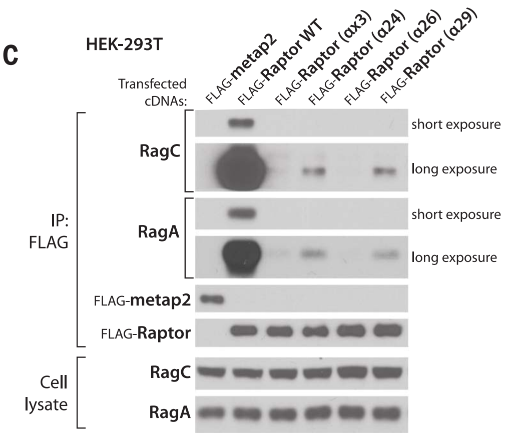
D

code	mutations in Rags	locations affected
RA1	Y31A	switch I
RA2	D35A, E46A	switch I, interswitch
RA3	A29F, D35Y, E46A	switch I, interswitch

B

code	mutations in Raptor	locations affected
αx3	A560F, R597E, D635A	helices α24, α26, α29
α24	N557A, A560F, E564A	helix α24
α26	C594A, R597Y, D598A	helix α26
α29	T634A, D635R, H636E	helix α29

C



E

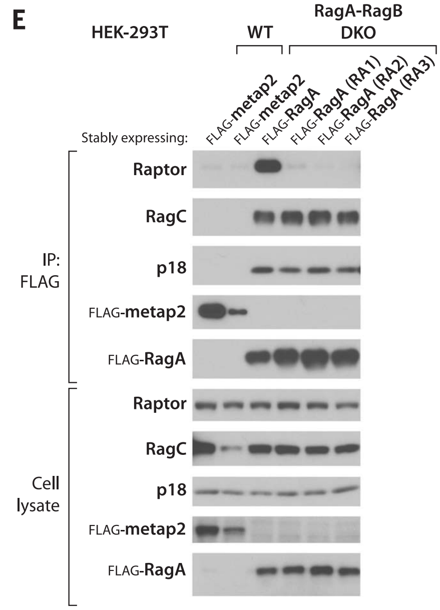


Fig. 2. Raptor-RagA interaction. (**A**) Three helices (α24, α26, α29) in the α-solenoid of Raptor directly engage the switch I face of GTP-loaded RagA. (**B**) Description of mutations introduced in the RagA-binding helices of Raptor. (**C**) Mutations in these helices render Raptor expressed via cDNA transient transfection unable to coimmunoprecipitate the endogenous RagA-RagC heterodimer. (**D**) Description of mutations introduced in the Raptor-binding region of RagA. (**E**) In RagA-RagB DKO HEK-293T cells, the transiently expressed RagA mutants in (D) cannot bind endogenous Raptor but have an intact capacity to bind to RagC and Ragulator, as assessed by its p18 subunit. Flag-metap2 was used as a control protein. Amino acid abbreviations: C, Cys; D, Asp; E, Glu; H, His; K, Lys; N, Asn; R, Arg; T, Thr; W, Trp; Y, Tyr.

RagA would rearrange and disrupt the binding sites for the three key Raptor helices (α24, α26, and α29), explaining why Raptor only interacts with the Rag heterodimer when RagA (or RagB) is GTP-bound (*8*).

The Raptor claw

The heterodimeric Rag GTPases exhibit rotational two-fold symmetry, such that their switch machineries are on opposite ends of the heterodimer and their nucleotide-binding pockets face the dimeric center (Fig. 3, A and B). Thus, if Raptor directly engages the switch I–containing face of RagA, how does it also detect the nucleotide state of RagC when its switch region is on the other side of the heterodimer?

Surprisingly, the refined map of the Raptor-Rag-Ragulator complex revealed a large stretch of electron density in the space between the two GTPase domains of the Rags. This density is acutely angled; it enters the space between the Rags from the Raptor side, loops back, and exits near the bound GDP on RagC. At first, it was not clear which protein contributed the density: RagA, RagC, Raptor, or a Ragulator component. By mining our particle images for rare Rag-Ragulator and Raptor-Rag subcomplexes and determining their structures to low resolution (see supplementary materials), we used a process of elimination to determine that the electron density belonged to Raptor (fig. S3A). We traced an amino acid backbone through the density, and the distinct positions of prolines assigned the side chains to amino acids 916 to 937 of Raptor (Fig. 3C). The electron density of this region is unaccounted for in the previous structure of Raptor (*31*), and considering its triangular shape, we named it the "Raptor claw." The primary sequence of the claw appears to be conserved in vertebrates and their closest invertebrate kin (fig. S3B).

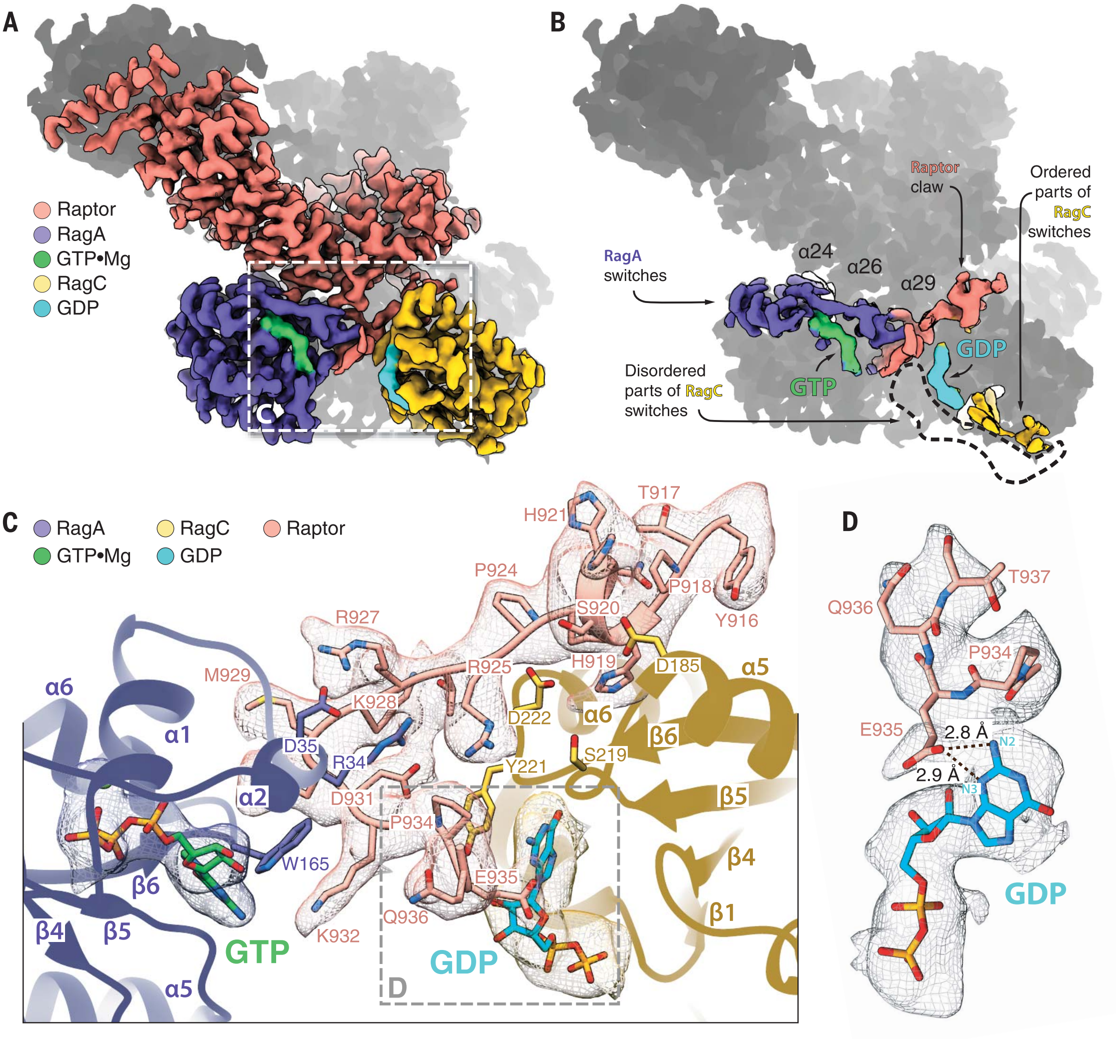


Fig. 3. Raptor-RagC interaction. (**A**) The binding interface between the Raptor α-solenoid and Rag GTPases. (**B**) The GTPase switch machineries of Rags are positioned at the opposite ends of the heterodimer. Note that in our structure, the switch machinery of RagC is largely disordered, and the circled region marks the space it would have occupied if it was fully resolved. (**C**) The Raptor claw is a 22–amino acid loop that enters the inter-Rag space depending on the nucleotide state of RagC. The claw forms interactions with both RagA and RagC, and its residue E935 engages RagC-loaded GDP directly. (**D**) The Glu^{935}-GDP interaction is driven by hydrogen bonding between the Glu^{935} side chain and N2 and N3 of the guanine of GDP. Amino acid abbreviations: D, Asp; E, Glu; H, His; K, Lys; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; W, Trp; Y, Tyr.

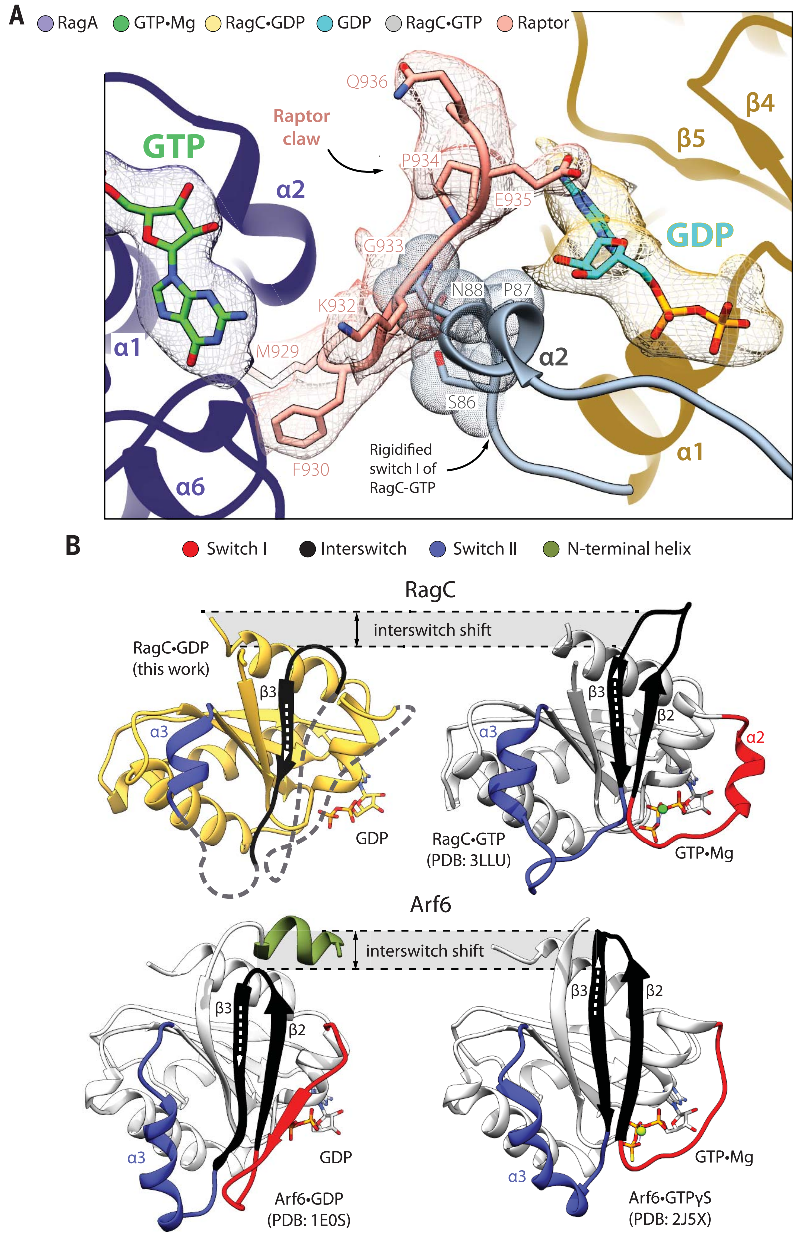
The N-terminal end of the claw contains a very short helical fragment α33a (amino acids 919 to 922) that anchors it in a groove between the GTPase (Asp^{185}, Asp^{186}, Lys^{188}) and CRD (Met^{279}) domains of RagC (fig. S3C). It enters the inter-GTPase domain space with a coil (Arg^{925}, Arg^{927}, Met^{929}) that first extends toward the CRD (Gln^{237}) and switch I (Ala^{33}, Arg^{34}, Asp^{35}) of RagA, and then loops back 130° and exits (Pro^{934}, Glu^{935}) near the nucleotide-binding site of RagC (Tyr^{221}) (Fig. 3C and fig. S3C). Interestingly, Glu^{935} of the claw is positioned close enough to the RagC-bound GDP to form hydrogen bonds with N2 and N3 of the guanine base (Fig. 3D). The switch machinery of RagC in the GDP-bound form in our cryo-EM structure is largely disordered, allowing us to model only parts of switch II and the interswitch. The remaining modules of the switch machinery appear dynamic.

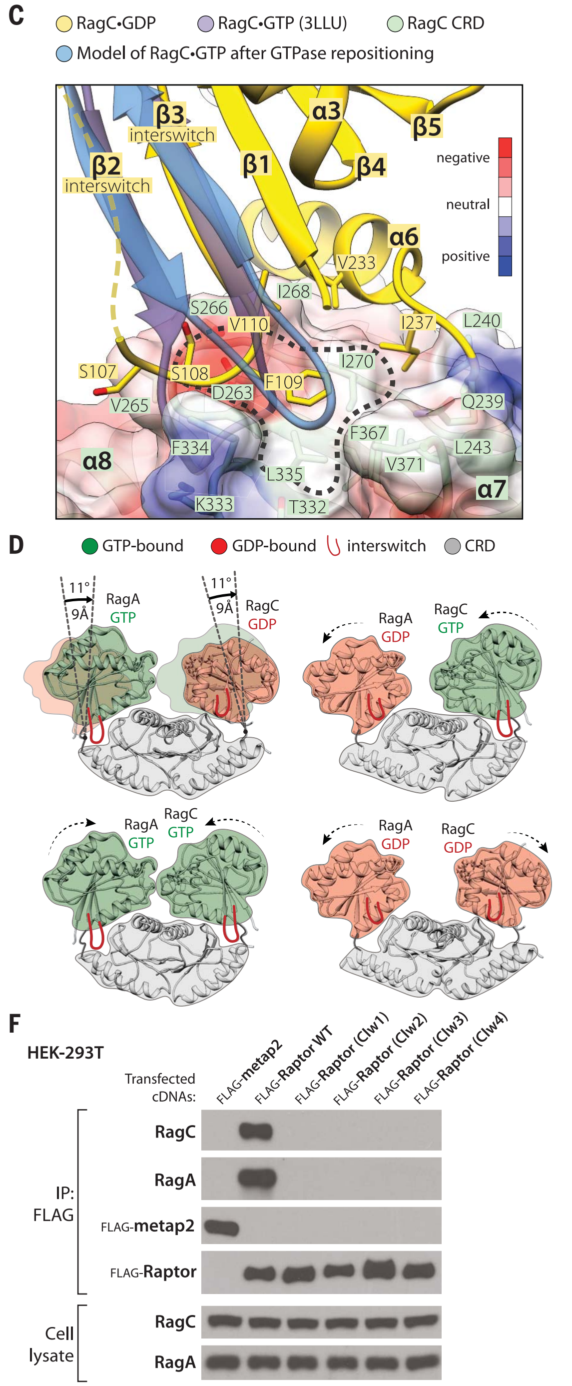
Detection by Raptor of the RagC nucleotide state

Because the Raptor claw can reach through the inter-Rag space to contact the switch machinery of RagC, we reasoned that it could serve as a nucleotide-state detector and explain why Raptor interacts with the Rag heterodimer when RagC (or RagD) is GDP-bound (*19*). To examine this possibility, we first superimposed a publicly available structure of the GTP-bound GTPase domain of RagC (PDB ID 3LLU) onto that of GDP-bound RagC from our cryo-EM structure. This analysis revealed that by causing switch I to rigidify, the loading of RagC with GTP would lead residues Ser^{86} and Asn^{88} of switch I to clash with the tip of the claw (residues Lys^{932} and Gly^{933}; see Fig. 4A).

Moreover, upon the GDP-to-GTP transition, the interswitch strand β3 of RagC shifts its register by two residues (Cα shift of 5.8 Å) in a manner similar to what has been observed in Arf GTPases (*32*) (Fig. 4B and movie S3). In GDP-bound RagC, Ser^{108} and Phe^{109} from the β2-β3 loop (107 to 110) of the interswitch rest on the side of a CRD pocket formed by four strands of the antiparallel β sheet and two bottom-face α helices (Fig. 4C). The two-residue register shift caused by GTP binding makes the interswitch stick outward, such that its β2-β3 loop would clash with the CRD pocket. To avoid such a clash, the interswitch loop would have to shift its position and engage with a larger area of the CRD pocket, leading the entire GTPase domain to reposition itself closer to the central axis of the Rag heterodimer (Fig. 4D). We modeled such domain repositionings in RagA and RagC in the context of the intact heterodimer because we observed no major domain rearrangements in the Rags upon Raptor binding, as evidenced by fitting the Rag heterodimer from the supercomplex into an 8.9 Å resolution map of the Rag-Ragulator subcomplex (fig. S3, A and D). This analysis showed that depending on the nature of the nucleotide, the space between the two GTPase domains becomes closed (RagA•GTP-RagC•GTP state), shifted toward the RagA (GDP-GTP) or RagC (GTP-GDP) side, or widened (GDP-GDP) (Fig. 4D).

This altered accessibility of the inter-Rag space likely is the basis of the ability of the claw to detect the nucleotide state of RagC. In the RagA•GTP-RagC•GDP state we captured,





E

code	mutations in Raptor	locations affected
Clw1	Δ(916-936)	Claw deleted
Clw2	H919A, S920A, H921A	Claw entrant helix
Clw3	R925A, R927A, K928A	Claw bend
Clw4	G933R, P934G, E935R	Claw GDP-proximal

Fig. 4. The dynamics of the Rag-Raptor interaction. (**A**) Loading of GTP by RagC would trigger its switch I to rigidify and clash with the Raptor claw. The structure of the GTP-loaded GTPase domain of RagC (PDB ID 3LLU) was superimposed with our cryo-EM structure of RagC•GDP. (**B**) The organization of the switches in RagC (top) and Arf6 (bottom) [PDB ID: 1E0S, 2J5X (*37*, *38*)]. Note that although the switches change their positions in GTP- versus GDP-loaded states, the core of the structure does not move. Even though a large proportion of the switches in our GDP-loaded RagC structure are disordered, we observed that the interswitch changes its register by two residues, in a manner similar to Arf GTPases. (**C**) The movement of the interswitch during the GDP-to-GTP transition would cause its loop to clash with the CRD pocket (circled area). Instead, the interswitch repositions itself such that it engages with the more central part of the CRD. The disordered interswitch strand β2 of RagC•GDP is drawn with a dashed line. The surface of the CRD is colored according to electrostatic potential (see the color key). (**D**) The shifting of the GTPase domains in the Rag heterodimer during GTP-GDP binding exchanges. GTPase domains loaded with GDP are positioned away from the central axis of the Rag heterodimer, and their interswitches are retracted. Loading of GTP causes the interswitch to extend and press the CRD pocket such that the entire GTPase domain becomes repositioned closer to the Rag central axis. The models were created by superimposing RagA•GTP with RagC•GDP through their CRDs (specifically by matching their β7 and α9). (**E**) Description of the Raptor claw mutants used in (F). (**F**) Elimination of the Raptor claw or mutations in its critical Rag-interacting regions prevent Raptor from coimmunoprecipitating the RagA-RagC heterodimer. Longer exposures of the RagA and RagC immunoblots did not reveal any signal for Raptor. Flag-metap2 was used as a negative control protein. Amino acid abbreviations: D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; S, Ser; T, Thr; V, Val.

the α-solenoid of Raptor engages switch I of RagA, and the claw penetrates the inter-Rag space because switch I of RagC is disordered (movie S4). In the RagA•GTP-RagC•GTP configuration, not only would switch I of RagC rigidify and push the claw away from the bound nucleotide, but also the entire GTPase-domain of RagC would shift to close the intersubunit space. The opposite state, RagA•GDP-RagC•GTP, appears completely incompatible with interacting with Raptor because RagA•GDP cannot support α-solenoid binding and RagC•GTP expels the claw. The primary sequences of the RagA and RagC switches matter. The α-solenoid of Raptor is able to associate with the switch

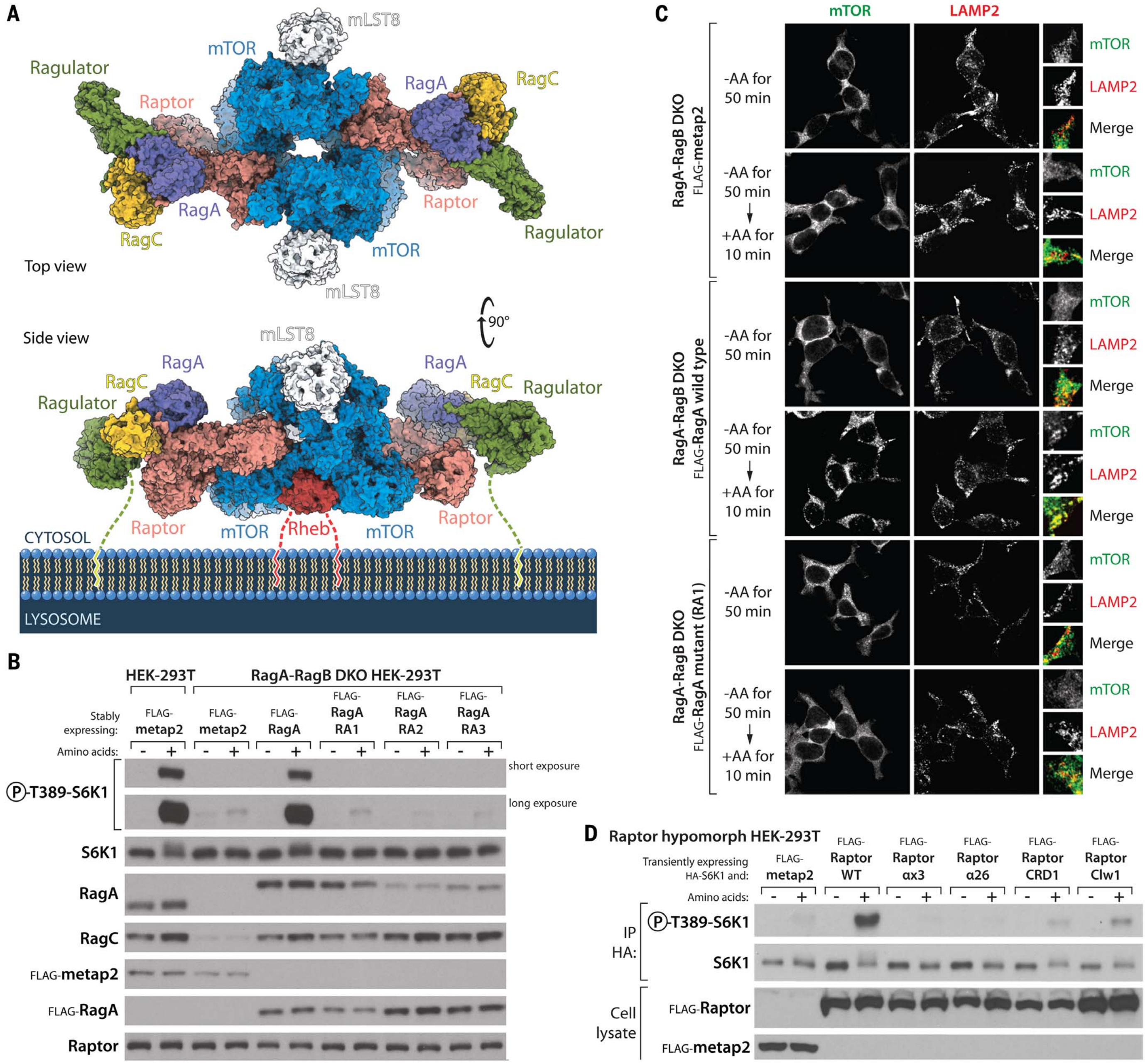


Fig. 5. mTORC1 on the lysosomal surface. (**A**) Model of a Rheb-activated mTORC1 dimer bound to two lysosomal-targeting Rag-Ragulator complexes. The Rag-Ragulator complex clamps mTORC1 to the surface of the lysosome. Note that both Ragulator and Rheb carry lipid-modified terminal residues that tether the supercomplex to the membrane. Our cryo-EM structure of the Raptor-Rag-Ragulator complex was superimposed with that of Rheb-bound mTORC1 [PDB ID 6BCU (*31*)]. (**B**) Stable expression in RagA-RagB DKO HEK-293T cells of RagA mutants that cannot bind Raptor (Fig. 2, D and E) does not restore mTORC1 activity, as assayed by the phosphorylation of S6K1 in amino acid–replete cells. Note that although the levels of the stably expressed wild-type and mutant Flag-RagA are relatively even as assessed with the anti-Flag antibody, the recognition of the RA2 and RA3 RagA mutants by the RagA antibody (monoclonal; CST D8B5) is reduced, which suggests that the mutated residues affect the epitope recognized by this antibody, which has not been disclosed. (**C**) Unlike expression of wild-type RagA, expression in RagA-RagB DKO HEK-293T cells of the Y31A RagA (RA1) mutant that cannot bind Raptor does not rescue the amino acid–induced colocalization of mTOR with lysosomes, as marked by the lysosomal protein LAMP2. (**D**) In cells hypomorphic for Raptor, transient expression of wild-type Raptor promotes S6K1 phosphorylation, whereas that of Raptor mutants defective in the RagA-Raptor interaction (αx3 and α26) does not. Mutants that cannot bind the CRD of RagC (CRD1) or lacking the claw (Clw1) activate S6K1 phosphorylation to a small degree. See supplementary materials and fig. S6B for a description of the Raptor hypomorph cell line.

machinery of RagA•GTP, but it would clash with the equivalent nucleotide state of RagC (fig. S3E). Technically, the open RagA•GDP-RagC•GDP state should allow the claw to make interactions within the inter-Rag space. However, threading of the claw involves not only RagC residues but also those from RagA. In the GDP-loaded state, the switch I residues of RagA, Arg^{34}, and Asp^{35} that contact the middle part of the claw (Fig. 3C) would become disordered. Ultimately, the entire RagA GTPase domain would shift away from the claw, leaving very little surface for any potential interactions. Beyond affecting the intersubunit space, the shifts in the GTPase domains we have modeled here are also likely involved in the cross-talk phenomenon that we previously observed in kinetic studies (*15*).

Given the extensive binding surface between RagA•GTP and the α-solenoid of Raptor, we considered that the claw might simply be a nucleotide detection mechanism for ensuring that RagC is GDP-bound and perhaps does not contribute to the strength of the Raptor-Rag interaction. Surprisingly, however, Raptor variants lacking the claw or with mutations in three critical regions—the entrant helix, the bend, and the part proximal to GDP (Fig. 4E)—could not bind the Rags (Fig. 4F) but still interacted with mTOR (fig. S3F).

In addition to the claw, the Raptor-Rag-Ragulator structure revealed a contact between the CRD of RagC and a loop (amino acids 795 to 806) extending from the α-solenoid of Raptor (fig. S4, A and B). This interface is largely driven by side chain–backbone hydrogen bonding (e.g., RagC Arg^{342} and Raptor Leu^{799}) and some backbone-backbone interactions (Raptor Gly^{801} and RagC Ser^{281}; fig. S4, C and D). Like the claw, this loop is also disordered in the previously reported structure of Raptor (*31*), and presumably its proximity to RagC rigidified it into a defined shape in our complexes. By mutating interfacing residues on Raptor, we found that deletion of residues 799 to 805 or a combination of three mutations (L799A, G801Y, F804D) weakened, but did not eliminate, the interaction of Raptor with the Rag GTPases without impacting that with mTOR (fig. S4E).

Collectively, our analyses of the role of interfacing residues between Raptor and the Rag heterodimer suggest a model in which Raptor first transiently binds to GTP-loaded RagA, forming interactions with its switch I region, and also weakly touches the CRD of RagC. Threading of the Raptor claw in between the GTPase domains of the Rags then generates a more stable interaction. The claw can fully engage only when RagC is loaded with GDP and its switch I is disordered. This would explain why RagC must be GDP-bound for the Rag heterodimer to interact with Raptor (*19*).

Model of mTORC1 docked on the lysosomal membrane

Translocation to the lysosomal surface is critical for the activation of mTORC1 because there it can interact with its kinase activator, the Rheb GTPase. A recent cryo-EM structure of mTORC1 bound to Rheb [PDB ID 6BCU (*31*)] provided a likely orientation of mTORC1 toward the lysosomal membrane because, like most small GTPases, Rheb associates with membranes through a C-terminal farnesyl group (*33*). Because both the Rheb-bound mTORC1 and Raptor-Rag-Ragulator structures contain Raptor, we superimposed them to generate a structural model of the Rheb-activated mTORC1 dimer tethered to the lysosome via Rag-Ragulator (Fig. 5A and movie S5).

In our original cartoons of mTORC1 on the lysosomal surface, we envisioned Rag-Ragulator as a pillar protruding from the lysosome with mTORC1 sitting atop it (*11*). This appears not to be the case. Instead, the Rag-Ragulator complex binds to the top and side of mTORC1, serving as a clamp that pushes mTORC1 down onto the lysosome (Fig. 5A). The p18/LAMTOR1 subunit of Ragulator has a 45–amino acid N-terminal tail with myristoyl and palmitoyl modifications on its end that is necessary for localizing Ragulator to the lysosome (*34*). Although we cannot resolve this tail in our structure, we predict that if extended, its length of ~14 nm would provide enough room to properly orient the assembled supercomplex of mTORC1-Rag-Ragulator on the lysosomal surface, as depicted in Fig. 5A. In fact, given the broad diameter range of lysosomes (0.1 to 1 μm) (*35*), the p18 tail should allow mTORC1 to make expansive searches of the lysosomal surface for the activated Rheb GTPase (fig. S5, A and B). Once two Raptor subunits of mTORC1 are clamped down by the Rag-Ragulator complexes, the search radius will be restricted. However, the resulting avidity effect of such dimeric binding will further improve the overall strength of the interaction with the lysosome and will facilitate the lasting activation of the mTOR kinase.

To study the consequences of eliminating the Raptor-Rag interaction on mTORC1 signaling, we took advantage of the HEK-293T RagA-RagB DKO cells lacking RagA and RagB (fig. S2C). These cells do not activate mTORC1 upon amino acid stimulation, as judged by the phosphorylation of S6K1, a canonical mTORC1 substrate. Although stable expression of wild-type RagA readily restored mTORC1 activity in amino acid–stimulated cells, that of the RagA mutants that cannot bind Raptor did not (Fig. 2, D and E, and Fig. 5B). Consistent with this finding, in the cells expressing RagA mutants deficient in Raptor binding, mTOR no longer localized with lysosomes upon stimulation with amino acids (Fig. 5C and fig. S6A).

It is not trivial to study the effects of Raptor mutations on mTORC1 signaling because Raptor is a cell-essential gene, and therefore it is not possible to generate cells lacking it. However, using CRISPR-Cas9, we were able to generate a HEK-293T cell line that has defective mTORC1 signaling because its Raptor alleles have partial loss-of-function mutations that also reduce their expression (fig. S6B) (see supplementary materials). Transient expression of wild-type Raptor in these cells activated mTORC1 signaling, whereas that of the Raptor mutants affecting the RagA binding site did not (Fig. 5D). The claw- and CRD-binding mutants were slightly less defective, as they provided a small fraction of the signal that wild-type Raptor did.

Intriguingly, a fairly abundant isoform of Raptor (isoform 3) that lacks the Rag-interacting α-solenoid (Δ504-661 compared to the canonical isoform 1) has been described (*36*), raising the possibility that dimeric mTORC1 complexes may exist containing two different Raptor isoforms, only one of which can bind to the Rag GTPases. Such mTORC1 heterodimers would likely associate more weakly with lysosomes, which may be useful in certain signaling contexts. mTORC1 dimers may also exist having only the non–Rag-binding Raptor isoform, which might play roles in cells (e.g., erythrocytes) that lack lysosomes. We attempted experiments with this isoform of Raptor but found that it expresses poorly in HEK-293T cells.

Our cryo-EM structure of Raptor, the defining subunit of mTORC1, in complex with Rag-Ragulator, allows us to model how mTORC1 would dock on the lysosomal surface, which is a key step in its activation. We have described the regulated interface between Raptor and Rag GTPases, the details of which are critical for understanding nutrient sensing at the molecular level and also for designing small molecules that are more specific to mTORC1 signaling than existing drugs such as rapamycin. These drugs also target the related protein complex mTORC2, which has other roles in the cell.

REFERENCES AND NOTES

1. R. A. Saxton, D. M. Sabatini, *Cell* **168**, 960–976 (2017).
2. D. M. Sabatini, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11818–11825 (2017).
3. J. Kim, K. L. Guan, *Nat. Cell Biol.* **21**, 63–71 (2019).
4. D. Mossmann, S. Park, M. N. Hall, *Nat. Rev. Cancer* **18**, 744–757 (2018).
5. A. J. Valvezan, B. D. Manning, *Nature Metab.* **1**, 321–333 (2019).
6. C. Buerger, B. DeVries, V. Stambolic, *Biochem. Biophys. Res. Commun.* **344**, 869–880 (2006).
7. J. Huang, B. D. Manning, *Biochem. J.* **412**, 179–190 (2008).
8. Y. Sancak *et al*., *Science* **320**, 1496–1501 (2008).
9. E. Kim, P. Goraksha-Hicks, L. Li, T. P. Neufeld, K. L. Guan, *Nat. Cell Biol.* **10**, 935–945 (2008).
10. A. Efeyan *et al*., *Nature* **493**, 679–683 (2013).
11. Y. Sancak *et al*., *Cell* **141**, 290–303 (2010).
12. L. Bar-Peled, L. D. Schweitzer, R. Zoncu, D. M. Sabatini, *Cell* **150**, 1196–1208 (2012).
13. N. Nakashima, E. Noguchi, T. Nishimoto, *Genetics* **152**, 853–867 (1999).
14. T. Sekiguchi, E. Hirose, N. Nakashima, M. Ii, T. Nishimoto, *J. Biol. Chem.* **276**, 7246–7257 (2001).
15. K. Shen, A. Choe, D. M. Sabatini, *Mol. Cell* **68**, 552–565.e8 (2017).

16. A. Wittinghofer, I. R. Vetter, *Annu. Rev. Biochem.* **80**, 943–971 (2011).
17. L. Bar-Peled *et al*., *Science* **340**, 1100–1106 (2013).
18. N. Panchaud, M. P. Péli-Gulli, C. De Virgilio, *Sci. Signal.* **6**, ra42 (2013).
19. Z. Y. Tsun *et al*., *Mol. Cell* **52**, 495–505 (2013).
20. C. S. Petit, A. Roczniak-Ferguson, S. M. Ferguson, *J. Cell Biol.* **202**, 1107–1122 (2013).
21. K. Shen, D. M. Sabatini, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9545–9550 (2018).
22. J. Okosun *et al*., *Nat. Genet.* **48**, 183–188 (2016).
23. T. Bepler *et al*., *Nat. Methods* 10.1038/s41592-019-0575-8 (2019).
24. T. Wagner *et al*., *Commun. Biol.* **2**, 218 (2019).
25. R. Yonehara *et al*., *Nat. Commun.* **8**, 1625 (2017).
26. M. E. G. de Araujo *et al*., *Science* **358**, 377–381 (2017).
27. K. Shen *et al*., *Nature* **556**, 64–69 (2018).
28. R. Gong *et al*., *Genes Dev.* **25**, 1668–1673 (2011).
29. J. H. Jeong *et al*., *J. Biol. Chem.* **287**, 29648–29653 (2012).
30. R. Nicastro, A. Sardu, N. Panchaud, C. De Virgilio, *Biomolecules* **7**, 48 (2017).
31. H. Yang *et al*., *Nature* **552**, 368–373 (2017).
32. S. Pasqualato, L. Renault, J. Cherfils, *EMBO Rep.* **3**, 1035–1041 (2002).
33. G. J. Clark *et al*., *J. Biol. Chem.* **272**, 10608–10615 (1997).
34. S. Nada *et al*., *EMBO J.* **28**, 477–489 (2009).
35. A. B. Novikoff, H. Beaufay, C. De Duve, *J. Cell Biol.* **2** (suppl.), 179–184 (1956).
36. C. Sun, C. Southard, A. Di Rienzo, *Mutat. Res.* **662**, 88–92 (2009).
37. J. Ménétrey, E. Macia, S. Pasqualato, M. Franco, J. Cherfils, *Nat. Struct. Biol.* **7**, 466–469 (2000).
38. S. Pasqualato, J. Ménétrey, M. Franco, J. Cherfils, *EMBO Rep.* **2**, 234–238 (2001).

ACKNOWLEDGMENTS

We thank all members of the Sabatini lab for helpful insights and suggestions, particularly K. Shen for help with the GTP-binding assays; the labs of T. Schwartz, C. Drennan, I. Cheeseman, J.-K. Weng, T. Baker, E. Fischer, and R. Williams, as well as the Broad Institute Center for the Development of Therapeutics and MIT Biology Structural Biology Core Facility, for advice, reagents, or access to instruments; the cryo-EM facilities of UMass Medical School, Brandeis University, and the Whitehead Institute, particularly K. Song (UMass), M. Rigney (Brandeis), and N. Watson (Whitehead), for training or access to electron microscopes; and the high-performance computing team at Whitehead (C. Andrew, P. McCabe, R. Taylor, and P. Macfarlane) for the installation and maintenance of data-processing servers. **Funding:** Supported by NIH grants R01 CA103866, R01 CA129105, and R37 AI47389, Department of Defense grant W81XWH-07-0448, and the Lustgarten Foundation (D.M.S.) and fellowships from the Tuberous Sclerosis Association (K.B.R.), the Koch Institute (X.G.), NIH F30 CA236179 (J.F.K.), the Charles A. King Trust (M.A.-R.), and a Saudi Aramco Ibn Khaldun Fellowship for Saudi Women (N.N.L.). D.M.S. is an investigator of the Howard Hughes Medical Institute and an American Cancer Society Research Professor. **Author contributions:** K.B.R. and D.M.S. designed the research plan and interpreted experimental results. K.B.R. performed all protein expression, purification, structural biology, and computational experiments with assistance from L.F.B., A.M.S.B., R.D., A.N., D.O., A.-C.P.F., S.X.Z., and D.L. X.G., J.F.K., M.A.-R., and N.N.L. designed and performed the signaling, localization, and protein-protein interaction experiments. E.J.B. assisted with cryo-EM data processing. K.B.R. wrote the manuscript and D.M.S. edited it. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Cryo-EM maps were deposited in the Electron Microscopy Data Bank under accession numbers EMD-20660 (Raptor-Rag-Ragulator) and EMD-20661 (Rag-Ragulator). The atomic model of Raptor-Rag-Ragulator was deposited in the Protein Data Bank under accession number 6U62. pFloat and pDarmo cloning plasmids were deposited on Addgene under accession numbers specified in table S1.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/468/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 and S2
Movies S1 to S5
References (*39–79*)

View/request a protocol for this paper from *Bio-protocol*.

13 May 2019; accepted 1 October 2019
Published online 10 October 2019
10.1126/science.aay0166

REPORT

FERROELECTRICS

Super-elastic ferroelectric single-crystal membrane with continuous electric dipole rotation

Guohua Dong[1]*, Suzhi Li[2]*, Mouteng Yao[1]*, Ziyao Zhou[1]†, Yong-Qiang Zhang[3], Xu Han[4], Zhenlin Luo[4], Junxiang Yao[5], Bin Peng[1], Zhongqiang Hu[1], Houbing Huang[6], Tingting Jia[5], Jiangyu Li[5], Wei Ren[1], Zuo-Guang Ye[7], Xiangdong Ding[2]†, Jun Sun[2], Ce-Wen Nan[8], Long-Qing Chen[9], Ju Li[10], Ming Liu[1]†

Ferroelectrics are usually inflexible oxides that undergo brittle deformation. We synthesized freestanding single-crystalline ferroelectric barium titanate ($BaTiO_3$) membranes with a damage-free lifting-off process. Our $BaTiO_3$ membranes can undergo a ~180° folding during an in situ bending test, demonstrating a super-elasticity and ultraflexibility. We found that the origin of the super-elasticity was from the dynamic evolution of ferroelectric nanodomains. High stresses modulate the energy landscape markedly and allow the dipoles to rotate continuously between the *a* and *c* nanodomains. A continuous transition zone is formed to accommodate the variant strain and avoid high mismatch stress that usually causes fracture. The phenomenon should be possible in other ferroelectrics systems through domain engineering. The ultraflexible epitaxial ferroelectric membranes could enable many applications such as flexible sensors, memories, and electronic skins.

Elasticity reflects the capability of a material to exhibit mechanical deformation and recovery. Elasticity and flexibility of material such as alloys and semiconductors are determined by the essential components (atoms and molecules, among others) and their interactions (chemical bonds) in the materials. In metals, elasticity originates from free transfer among electron components without changing the Fermi level substantially during mechanic deformation (*1*). Some alloys called shape-memory alloys can even achieve the shape recoverability with strain exceeding the elastic limit through the martesitic phase transition, leading to the super-elasticity (*2*). By contrast, the ionic or covalence bonds in oxides and semiconductors have much greater interaction strength but lack sufficient slip systems for dislocation movement under deformation, leading to a brittle fracture. Nevertheless, the phase transitions in nanostructural oxides and semiconductors with reduction in dimensionality (*3–5*) offer atomic displacement tolerance in coexisted phases (*6*) and trigger additional strain that exceeds the elastic limit of bulk materials and has the ability for recovery.

Ferroelectrics (FEs), as a class of functional oxides, have found broad applications due to the existence of switchable spontaneous polarization and its coupling to mechanic deformation. Flexible FE thin films and membranes are needed for advanced flexible FE electronics, which will fulfill the growing demands for applications (*7–9*). In recent years, novel peeling and transferring technologies developed for oxide thin films have provided a step forward toward fabricating high-quality epitaxial membranes (*10–15*). However, little has been done in investigating the presence, origin, and limit of elasticity and flexibility in FE thin films and membranes. This is an important topic pertinent to various research areas such as flexoelectricity and multiferroics, in which strong electrical-mechanical coupling effects can be generated through large bending deformation. Although the ionic or covalence bonds in FEs constrain the elasticity in general, the FE domain evolution could provide an alternative avenue to complete the super-elasticity in oxides instead of phase transformations. For example, perovskite $BaTiO_3$ (BTO) films (tetragonal phase), a classic FE and piezoelectric material, usually contain both *a* and *c* domains (*16*, *17*). The transition between these *a* and *c* domains may enable shape tolerability during the bending process.

We synthesized BTO/$Sr_3Al_2O_6$ (SAO) heterostructures by means of pulsed-laser deposition on (001) $SrTiO_3$ (STO) substrates (*18*). The samples were immersed in the deionized water to dissolve the sacrificial SAO layer to obtain freestanding BTO membranes (*10*). The synchrotron-based reciprocal space mapping (RSM) around the pseudocubic (002) and (103) reflections confirms the heteroepitaxial growth of the BTO (60 nm)/SAO film (Fig. 1B). We confirmed the *c* axis–oriented BTO/SAO heterostructure with the θ to 2θ pattern (fig. S1). We found the BTO showed a tetragonal phase with c/a = 1.024 and possessed along an out-of-plane polarization. The RSM pattern around the (002) and (103) peaks presented single-crystal characteristics in flexible BTO/polydimethylsiloxane (PDMS). We recorded a wide range of thicknesses for the RSM of films (figs. S2 and S3). We used scanning electron microscopy (SEM) and atomic force microscopy (AFM) to determine the surface morphology and roughness of the large-scale transferred BTO membrane (fig. S4).

We transferred large-area freestanding single-crystal BTO membranes onto a Cu grid with lacey carbon film. We imaged the 50-nm BTO membrane with a low-magnification plan-view scanning transmission electron microscopy (STEM) and atomic resolution high angle annular dark field (HAADF) STEM image (Fig. 1G). We also imaged freestanding and cross-sectional membranes (figs. S5 to S7). Under the irradiation of an electron beam, the freestanding BTO membrane looks like a translucent plastic layer on the porous carbon frame. This Z-contrast image presents a typical perovskite structure with four high-intensity Ba atomic columns and one low-intensity Ti atomic column (*19*). The precise atomic stacking matches well with the crystal structure of BTO. Interestingly, we identified the exquisite fringes we observed as equal inclination fringes with the so-called bend contours (*20*). The presence of exquisite bend contours in the hollow areas also proves the single-crystalline nature of the BTO membrane with a uniform thickness and the formation of the buckling state.

We then transferred the BTO membrane onto the Si substrate and tested the electromechanical properties using piezoresponse force microscopy (PFM). We applied V = ±10 V to write an out-of-plane FE domain–switching image (Fig. 1H). We also measured the polarization-voltage loops that the BTO membranes transferred on Pt/Si (fig. S8). We obtained explicit reversible switching of FE domains after positive- and negative-bias poling. We obtained the complete FE switching, as exhibited by the piezoelectric response loops (Fig. 1I).

During the preparation, some of the freestanding BTO membranes were folded (fig. S9). All these folded BTO membranes maintained

[1]Electronic Materials Research Laboratory, Key Laboratory of the Ministry of Education, School of Electronic and Information Engineering, State Key Laboratory for Mechanical Behavior of Materials, International Joint Laboratory for Micro/Nano Manufacture and Measurement Technology, Xi'an Jiaotong University, Xi'an 710049, China. [2]State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China. [3]Center for Advancing Materials Performance from the Nanoscale (CAMP-Nano) and Hysitron Applied Research Center in China (HARCC), State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China. [4]National Synchrotron Radiation Laboratory and CAS Key Laboratory of Materials for Energy Conversion, Department of Physics, University of Science and Technology of China, Hefei 230026, China. [5]Shenzhen Key Laboratory of Nanobiomechanics, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. [6]Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing 100081, China. [7]Department of Chemistry and 4D LABS, Simon Fraser University, Burnaby, BC V5A 1S6, Canada. [8]State Key Lab of New Ceramics and Fine Processing, School of Materials Science and Engineering, Tsinghua University, Beijing 100084, China. [9]Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA 16802, USA. [10]Department of Nuclear Science and Engineering and Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: mingliu@xjtu.edu.cn (M.L.); ziyaozhou@xjtu.edu.cn (Z.Z.); dingxd@mail.xjtu.edu.cn (X.D.)

integrity and continuity without any cracks. These films behaved like flexible metallic films instead of brittle oxides. Furthermore, we bent the freestanding BTO membranes into several multilayer rolls (fig. S10), which were cut with a focus ion beam (FIB) to prepare cross-sectional transmission electron microscopy (TEM) samples (Fig. 2A). We investigated the structure by choosing four positions with different bending strains labeled with W, X, Y, and Z (Fig. 2A). These areas had different curvatures.

The Y area possessed the strain of up to 3.1% in the BTO roll (Fig. 2A and fig. S11). On the bottom of the BTO rolls, we can distinguish the five BTO layers, and we collected the selected-area electron diffraction pattern at the marked area covering all the layers (fig. S12). We found that all of the layers had the same orientation. The atomic-scale STEM image of the interface between two BTO layers in the Z area indicated that the connection between these BTO layers is by physical adsorption instead of chemical bonding (Fig. 2A). We found a clear Ba atomic column arrangement with a fan-shaped distribution and a continuous strain gradient from the cross-section atomic-scale STEM image of a bent BTO membrane (Fig. 2B). We distinguished two different strained regions, the tensile zone on the top and the compressive strain zone on the bottom. We mapped the accurate strain (ε) distribution directly at the atomic scale by using geometrical phase analysis (*19*, *21*). We calculated strain maps of ε_{xx} and ε_{yy}, which range from –5 to +5% (Fig. 2C).

When an FE membrane is bent, the tensile and compressive regions undergo different local polarization changes at the atomic scale (*22*, *23*). We used HAADF-STEM to image a film with an overlay of the spontaneous polarization direction of BTO (Fig. 2D). We observed multiple polarization orientations in the region where the maximum strain is achieved ε_{max} – 3.1%. Compared with the RSM results, the domain structure was transformed from a single *c*-domain state to a multidomain state by the strain gradient during bending. We showed a more explicit relationship between the lattice distortion and the associated polarization using magnified images from different regions demarcated in Fig. 2, D and E, which is associated with the flexoelectricity (*24*, *25*). We analyzed a series of STEM images with varying conditions of strain to explore the variation of polarization orientation (fig. S13). With the increase of bending stress, the domain state evolved from *c* domains to *a* domains, and we found an intermediate state. The dipoles rotate continuously between the *a* and *c* domains by means of the intermediate state, which does not match the domain characteristics of a tetragonal FE phase. To understand the influence of strain on polarization, we performed atomistic simulations in bent BTO membranes (fig. S14). The simulations that were used had a membrane thickness of 20 nm with the ε_{max} (close to 3.1%) applied. We selected four typical areas to examine the local dipole configuration (Fig. 2F). The polarization orientation was correlated with the strain states in the simulation. Typically, the polarization oriented upward in the compressive strain and tilted downward in the tensile strain while becoming horizontal near the neutral strain region, which was consistent with the experimental observations.

The formation of rolls implies that the freestanding BTO membranes had excellent flexibility, which is a rare feature for functional oxide materials. We manipulated BTO nanobelts using in situ SEM. We fabricated freestanding BTO nanobelts using the FIB and handled them with the nano-manipulator tip in the SEM (fig. S15). We carried out the bending tests on BTO nanobelts with thicknesses of 60 and 120 nm, respectively (Fig. 3). We used one nano-manipulator tip to hold the BTO nanobelt (120 nm), and we used another nano-manipulator tip to push the BTO nanobelt to bend. This process allowed us to bend BTO nanobelts into different curvatures (Fig. 3A). We achieved a maximum strain of ~10% in the BTO nanobelt without any fracture (fig. S16), demonstrating excellent flexibility in this oxide material. We measured the evolution of the elastic properties of a BTO nanobelt (60 nm in thickness) at a bending angle of 40° with different holding durations (Fig. 3B). We were able to instantly recover the strain from this BTO nanobelt. We recovered a large portion of the strain instantaneously even when increasing the bending angle up to 80° (Fig. 3C and movies S1 to S3).

The conventionally brittle BTO has good shape recovery when prepared as nanobelts. The super-elasticity usually is associated with reversible phase transformations like those seen in shape-memory alloys (*2*, *26*). The physical nature of elasticity, in general, is the tolerance that allows large shape deformation and the energy conversion that provides the ability of recovery. We examined the BTO domain architectures under different bending configurations through TEM bright-field imaging (fig. S17), which confirmed that the domain structure is highly dependent on strain state. Additionally, we compared the Raman spectra of the flat and the rolled BTO membranes on glass substrates, which allowed us to exclude structural phase transition during the bending process (fig. S18). In the freestanding BTO membranes, we believe the FE domain switching to be responsible for the observed flexibility and super-elasticity.

We carried out atomistic simulations on the microstructure evolution under mechanical bending in the BTO membranes at 300 K. Figure 4A shows the variation of bending

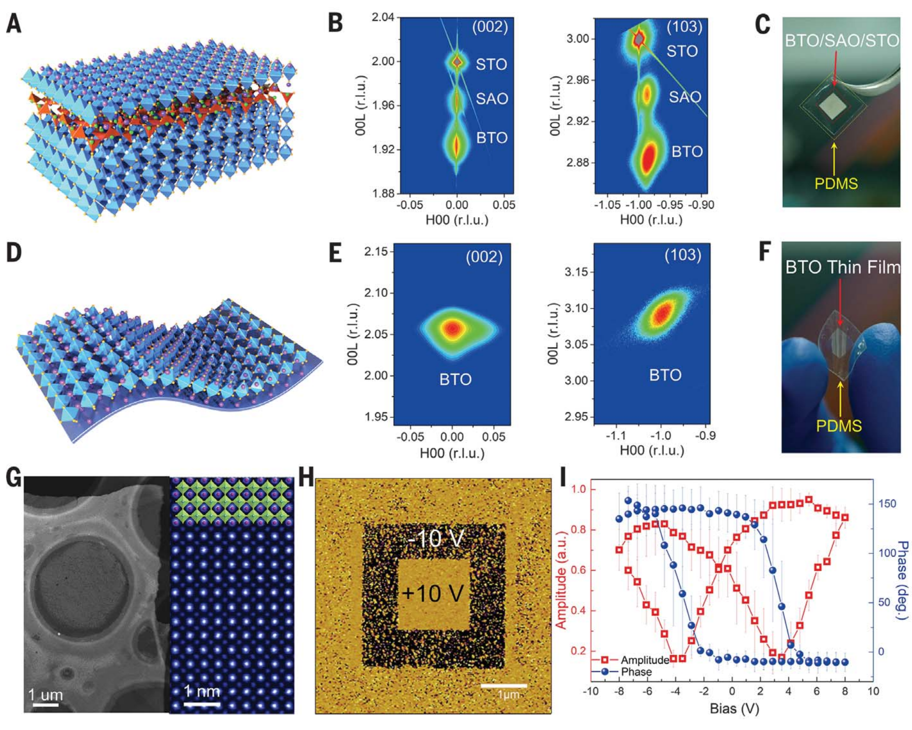


Fig. 1. Synthesis and characterizations of freestanding BTO membranes. Schematics (**A** and **D**), reciprocal space maps (**B** and **E**) around the (002)- and (103)-diffraction condition of STO, and optical photographs (**C** and **F**) of the BTO/SAO/STO heterostructure and flexible BTO membrane, respectively. r.l.u., relative light units. (**G**) Plane-viewed HAADF-STEM images of freestanding BTO membrane supported by a Cu grid with lacey carbon film. (**H**) Out-of-plane PFM phase image of BTO transferred on Pt/Si. (**I**) Piezoelectric response of the BTO membranes transferred on Pt/Si. a.u., arbitrary units.

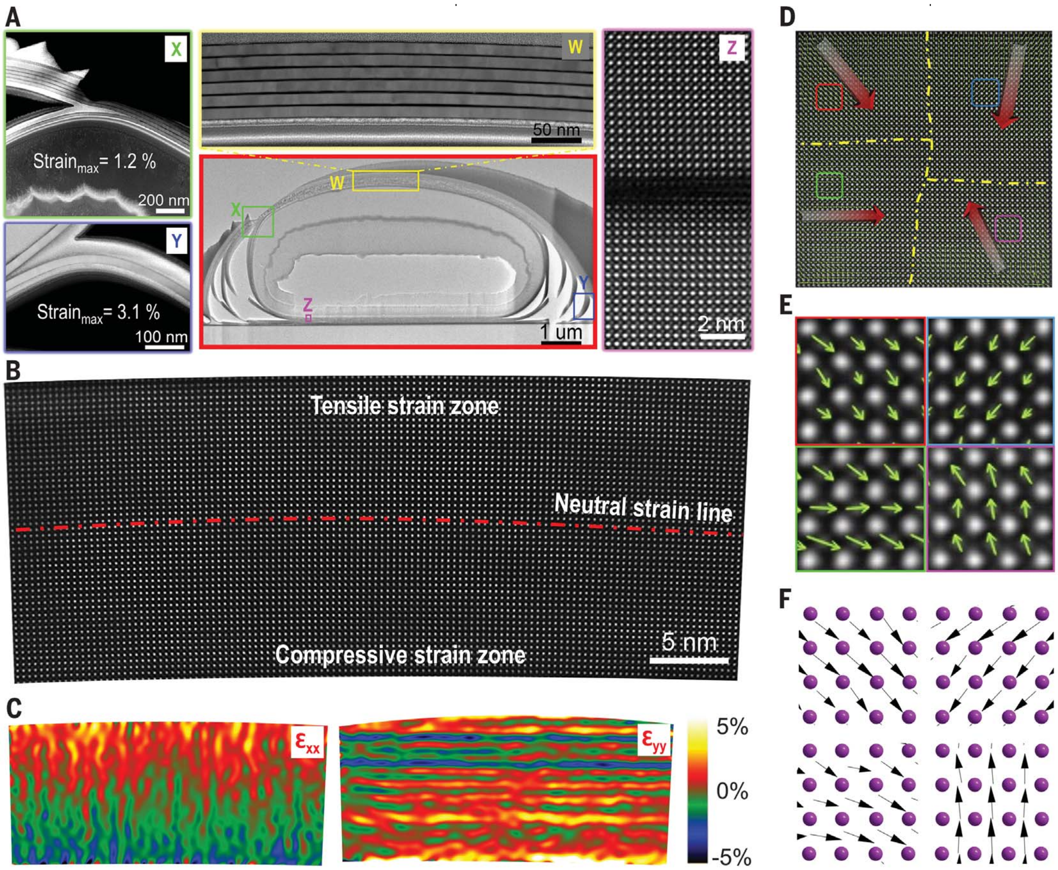


Fig. 2. Atomic-scale spatial distribution of continuous and curved lattice. (**A**) Cross-sectional STEM and TEM images of multilayer BTO roll with magnified images of the different regions, marked by W, X, Y, and Z. W, cross-sectional STEM image of freestanding BTO multilayer; X, the region with strain$_{max}$ ~ 1.2%; Y, the region with strain$_{max}$ ~ 3.1%; Z, the region with near-zero strain. (**B** and **C**) Atomic-scale STEM and geometric phase analysis images of BTO. (**D**) Atomic-scale STEM together with Ti displacement of the region with strain$_{max}$ ~ 3.1%. (**E**) Magnified images of the four regions from (D). Dipole displacements are amplified by a factor of 15 for clarity. (**F**) The dipole configurations in four selected areas of a bent membrane by atomistic simulations. Dipole displacements are amplified by a factor of 50 for clarity.

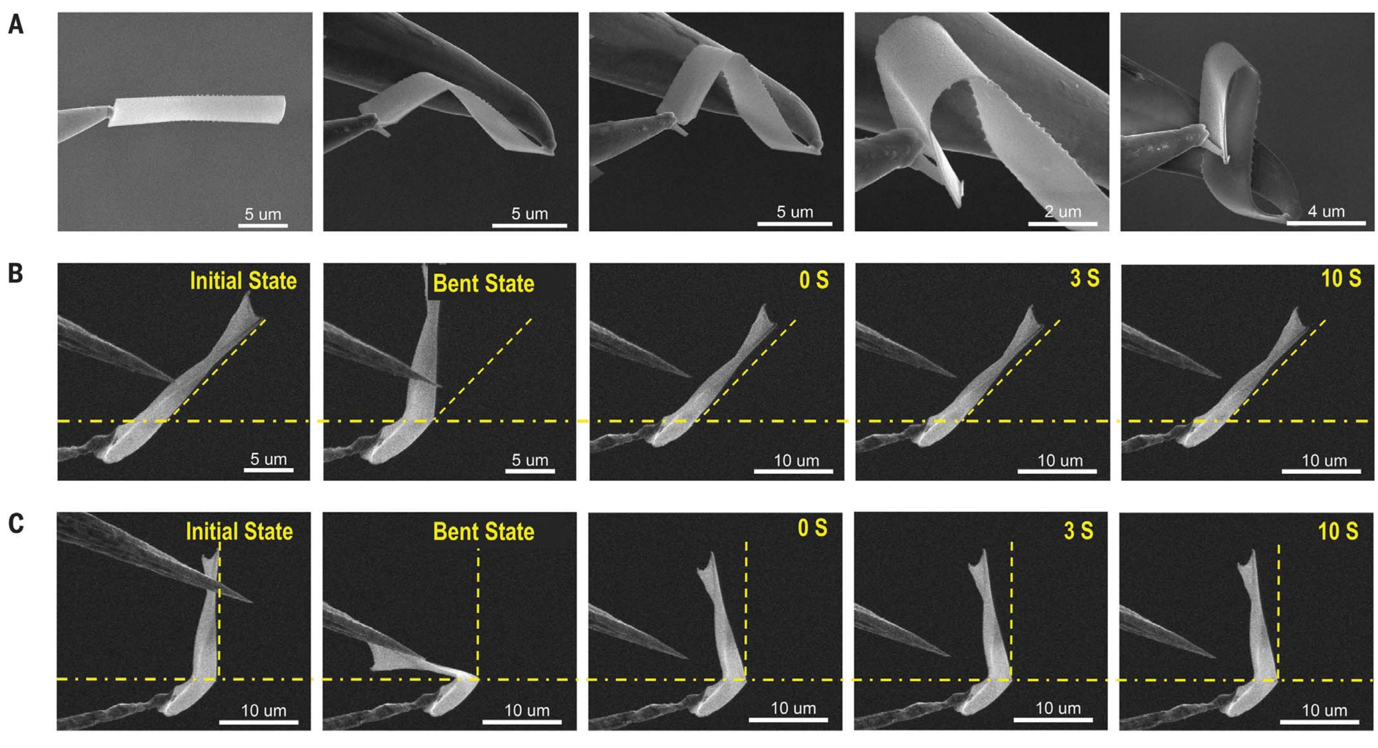


Fig. 3. In situ SEM bending test of BTO nanobelts. (**A**) Series of SEM images with the bending process of a BTO nanobelt (20 μm by 4 μm by 120 nm). (**B** and **C**) SEM images of the recovery process after the different bending states for a BTO nanobelt (20 μm by 4 μm by 60 nm).

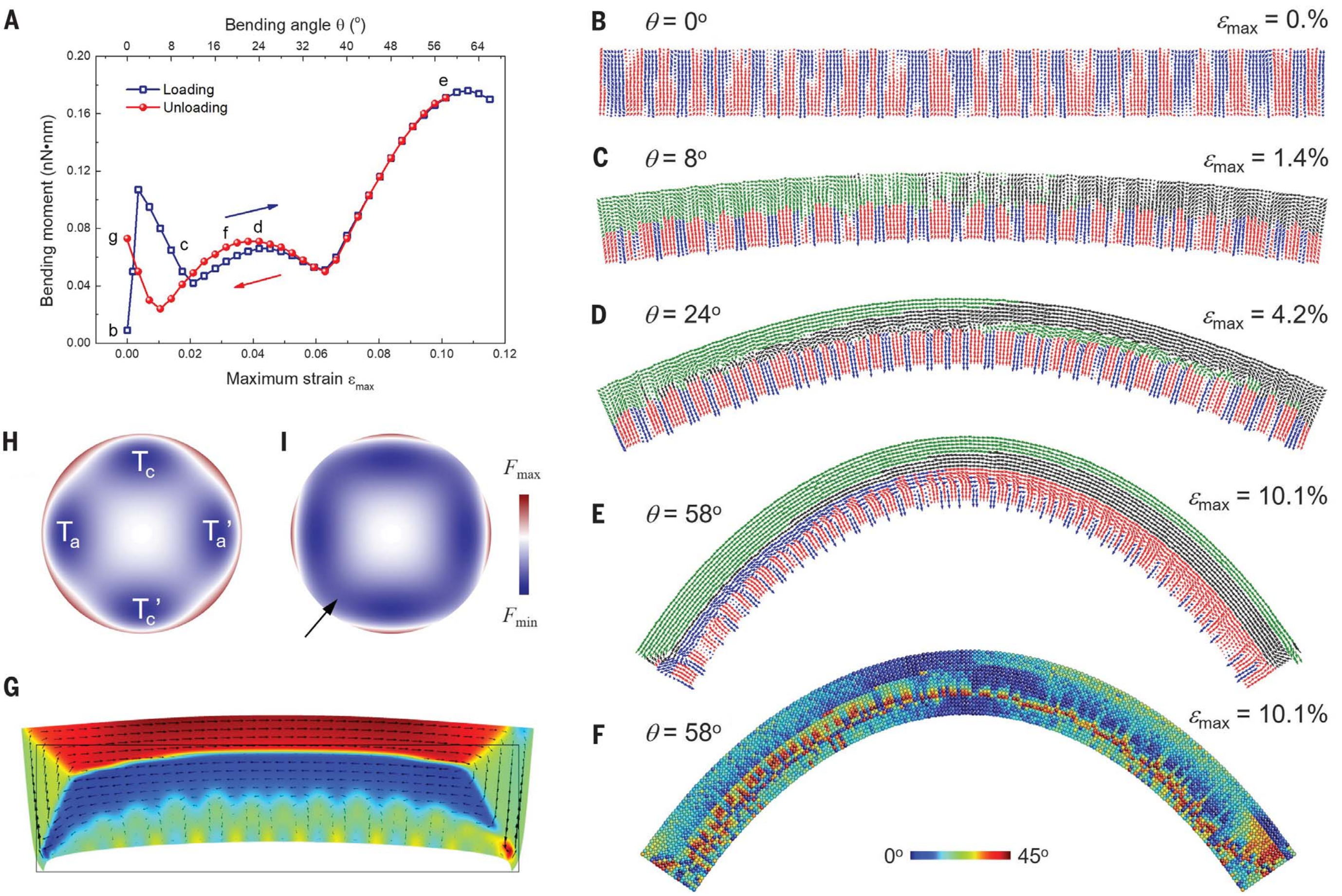


Fig. 4. Shape recovery of the BTO membrane upon bending at 300 K. (**A**) Variation of bending moment with bending angle θ (and maximum strain ε_{max}) upon loading and unloading with sample width of 8 nm. (**B** to **E**) Typical snapshots of dipole configurations at different bending angles. The black and green arrows represent the dipoles with *x* components left and right, respectively, in the tensile region. The red and blue arrows represent the dipoles with *y* components up and down, respectively, in the compressive region. Dipole displacements are amplified by a factor of 50 for clarity. (**F**) Spatial distribution of angle deviation η with respect to the *a* or *c* nanodomain at θ = 58°. η characterizes the deviation of dipole rotation with respect to the *a* or *c* domains by using the parameter $\eta = \pi/4 - |\varphi - (n^*\pi/2 + \pi/4)|$ for $\varphi \in [n^*\pi/2, (n+1)^*\pi/2]$, where φ is the rotation angle of polarization concerning the central axis in each unit cell and *n* runs from 0 to 3. (**G**) Phase-field modeling of the domain structures of the bent BTO membrane. The colors are coded according to the orientations of the domains. (**H** and **I**) Schematic illustration of free energy landscape of bulk BTO and nanomembrane under bending, respectively. (H) T_a (T_a') and T_c (T_c') refer to the tetragonal phases with *a* domains (polarized leftwards and rightwards) and *c* domains (polarized upwards and downwards) in bulk. (I) The energy barrier between *a* and *c* nanodomains drops markedly (indicated by the black arrow), allowing continuous rotation of dipoles in the bent membrane.

moment as a function of bending angle θ (and maximum strain ε_{max}) upon a loading and unloading cycle. The membrane deformed elastically, followed by the plastic flow. Further bending induced a continuous increase of the moment until the final drop, corresponding to the crack nucleation. Upon unloading before fracture, it almost traced back, resulting in a typical super-elastic behavior. The total recoverable strain could reach ~10%, close to our experimental value. We calculated the loss factor of the mechanical hysteresis to be only ~0.02, far below that of ~0.14 in zirconia shape-memory ceramics (*6*) and ~0.2 in Ti–Ni–Cu shape-memory alloys (*27*) at small scale.

To understand the mechanism responsible for such a substantial shape recovery, we examined the evolution of dipole configurations at the atomic scale. Before loading, we identified the lattices as *c* nanodomains with polarization either upward (red arrows) or downward (blue arrows) (Fig. 4B). In the tensile regions, the vertical dipoles began to rotate clockwise or anticlockwise locally at a smaller bending angle (Fig. 4C). Large bending can promote the local nanodomains to merge and transform into *a* nanodomains with polarizations oriented left (black arrows) or right (green arrows) in the tensile region (Fig. 4D). As we applied the considerable bending angle, one oriented *a* nanodomain grew to cover the whole free surface, forming a 180° domain boundary with a few steps (Fig. 4E). The polarizations prefer to be upward in the compressive region such that the 180° domain switching occurs from downward to upward in some local areas. Because the twin wall energy is usually very small in FEs (*28*–*31*), the energy cost of forming such high-density nanodomains is not that high. The dipoles almost switch back upon unloading (movies S4 and S5). However, the *a* domains cannot revert to *c* domains completely in the tensile region (fig. S19). Thus, the membrane cannot achieve a full shape recovery, because of the irreversible transformation of domain patterns locally.

We know the bulk BTO is brittle under a minimal strain. The fracture is usually originated from the initiation of crack near the internal flaw or along the grain boundaries. Two reasons might account for the flexibility in our BTO membranes. First, the flaw density is markedly reduced in such a smaller sample in comparison with its bulk counterpart. The paucity of flaw would avoid stress concentration and suppress the nucleation of crack in the interior of the sample. Unlike the bulk where

180° and 90° domains usually form, we observed that dipoles rotate continuously in a transition zone connecting *a* and *c* domains (Fig. 4F and fig. S20), similar to the rotation of polarizations across the 180° twin walls observed in the tetragonal $PbZr_{0.2}Ti_{0.8}O_3$ films experimentally (*32*, *33*). The flexoelectric effect is the primary driving force for the polarization rotation in *a* and *c* nanodomains far away from the neutral strain line (*34*) but only contributes partially in the transition zone (supplementary materials). The formation of such a continuous transition zone could largely eliminate the mismatch stress in the coexisting *c* and *a* nanodomains at high strain, avoiding the mechanical failure by the sharp domain switching. This transition zone exists in samples with a thickness ranging from 6.4 to 11.2 nm (fig. S21). In addition, Fig. 4G shows the domain structure of the bent BTO membrane under a high strain, obtained from phase-field simulation. We observed a similar dipole configuration, consistent with the experimental and the atomistic simulation results.

We emphasized that the continuous transition occurs when the size reduces to the nanoscale (*35*–*37*). According to the phenomenological Landau theory (*38*, *39*), four minima in the energy landscape exist in the bulk stress-free state, corresponding to two oriented *a* domains and *c* domains separately (Fig. 4H). Following the smaller-is-stronger trend (fig. S21), a high stress could be reached at nanoscale, and such high stress could modify the landscape markedly. Consequently, the energy barrier between *a* and *c* domains would be reduced, generating the metastable states in which dipoles are able to stay with some deviations from *a* or *c* domains (Fig. 4I). This metastable state in the transition zone could easily revert to the *a*- or *c*-domain state when removing the load. This would account for why we might observe shape recoverability at nanoscale rather than in bulk.

Combined with the intrinsic piezoelectricity and ferroelectricity and the flexoelectric effects (fig. S22 and S23) integrated by lattice and strain, especially in a two-dimensional structure, the reported super-elasticity in the BTO membranes is expected to provide more degrees of freedom for a wide range of applications in nanotechnology. Additionally, the flexible FE membranes can also serve as a viable platform to explore strain-triggered correlated phenomena, such as functionality enhancement, FE domain engineering, and phase transitions, in future research.

REFERENCES AND NOTES

1. T. B. M. C. Barret, *Structure of Metals*, (Pergamon, Oxford, UK, 1992).
2. J. F. Gómez-Cortés *et al*., *Nat. Nanotechnol.* **12**, 790–796 (2017).
3. G. Cheng *et al*., *Nat. Nanotechnol.* **10**, 687–691 (2015).
4. S. L. Wang, Z. W. Shan, H. Huang, *Adv. Sci.* **4**, 24 (2017).
5. J. Cao *et al*., *Nat. Nanotechnol.* **4**, 732–737 (2009).
6. A. Lai, Z. Du, C. L. Gan, C. A. Schuh, *Science* **341**, 1505–1508 (2013).
7. T. Leydecker *et al*., *Nat. Nanotechnol.* **11**, 769–775 (2016).
8. F. R. Fan, W. Tang, Z. L. Wang, *Adv. Mater.* **28**, 4283–4305 (2016).
9. Q. Zheng, B. J. Shi, Z. Li, Z. L. Wang, *Adv. Sci.* **4**, 23 (2017).
10. D. Lu *et al*., *Nat. Mater.* **15**, 1255–1260 (2016).
11. S. R. Bakaul *et al*., *Nat. Commun.* **7**, 10547 (2016).
12. L. Chang, L. You, J. L. Wang, *Jpn. J. Appl. Phys.* **57**, 7 (2018).
13. D. Ji *et al*., *Nature* **570**, 87–90 (2019).
14. Y. Qi *et al*., *Nano Lett.* **10**, 524–528 (2010).
15. Z. G. Zuo *et al*., *J. Phys. D Appl. Phys.* **45**, 5 (2012).
16. L. Hong *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **77**, 7 (2008).
17. S. Sahoo *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **76**, 092108 (2007).
18. Materials and methods are available as supplementary materials.
19. C. Dubourdieu *et al*., *Nat. Nanotechnol.* **8**, 748–754 (2013).
20. E. Spiecker, *Ultramicroscopy* **92**, 111–132 (2002).
21. J. L. Rouvière, E. Sarigiannidou, *Ultramicroscopy* **106**, 1–17 (2005).
22. I. Stolichnov *et al*., *Nano Lett.* **15**, 8049–8055 (2015).
23. Z. B. Chen *et al*., *Phys. Rev. Lett.* **118**, 7 (2017).
24. T. D. Nguyen, S. Mao, Y. W. Yeh, P. K. Purohit, M. C. McAlpine, *Adv. Mater.* **25**, 946–974 (2013).
25. W. J. Chen, Y. Zheng, X. Feng, B. Wang, *J. Mech. Phys. Solids* **79**, 108–133 (2015).
26. Y. Tanaka *et al*., *Science* **327**, 1488–1490 (2010).
27. J. San Juan, M. L. Nó, C. A. Schuh, *Nat. Nanotechnol.* **4**, 415–419 (2009).
28. D. D. Viehland, E. K. H. Salje, *Adv. Phys.* **63**, 267–326 (2014).
29. Y. M. Jin, Y. U. Wang, A. G. Khachaturyan, J. F. Li, D. Viehland, *J. Appl. Phys.* **94**, 3629–3640 (2003).
30. D. Viehland, *J. Appl. Phys.* **88**, 4794–4806 (2000).
31. Y. M. Jin, Y. U. Wang, A. G. Khachaturyan, J. F. Li, D. Viehland, *Phys. Rev. Lett.* **91**, 4 (2003).
32. G. De Luca *et al*., *Adv. Mater.* **29**, 1605145 (2017).
33. S. Cherifi-Hertel *et al*., *Nat. Commun.* **8**, 15768 (2017).
34. G. Catalan *et al*., *Nat. Mater.* **10**, 963–967 (2011).
35. D. A. Cogswell, M. Z. Bazant, *Electrochem. Commun.* **95**, 33–37 (2018).
36. Z. Zhang *et al*., *Phys. Rev. Lett.* **111**, 5 (2013).
37. D. A. Cogswell, M. Z. Bazant, *Nano Lett.* **13**, 3036–3041 (2013).
38. Y. J. Gu *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **89**, 11 (2014).
39. Y. L. Li, L. Q. Chen, *Appl. Phys. Lett.* **88**, 3 (2006).

ACKNOWLEDGMENTS

We thank D. Wang, H. Nan, and L. Lu (Xi'an Jiaotong University) for developing the APPSA software for locating the atomic positions and TEM sample preparation and R. Ramesh (University of California, Berkeley), X. Wang, X. Zhuang (Tongji University), and C. Guo (Beijing Institute of Technology) for fruitful discussions. **Funding:** We acknowledge the support from the National Key R&D Program of China (grant 2018YFB 0407601), the National Science Foundation of China (grants 11534015, 51602244, 51802248, 11574246, 51931004, 51621603, 51788104, 11374010, 11675179), the National 111 Project of China (grant B14040), the National Key R&D program (2016YFA0300102), National Key Research and Development Program of China (grant 2016YFA0201001 to J.Y., T.J., and J.L.), and the Natural Sciences and Engineering Research Council of Canada (grant 203773 to Z.-G.Y.). **Author contributions:** M.L. and Z.Z. designed and supervised the project. G.D. and M.Y. fabricated the films. G.D. conducted STEM and TEM experiments and the data analysis. S.L., J.S., Ju.L., and X.D. carried out the atomistic simulations. Y.-Q.Z. performed the in situ SEM experiments. X.H. and Z.L. conducted the RSM tests. J.Y., T.J., and Ji.L. conducted the PFM tests. H.H. and L.-Q.C. performed the phase-field simulations. B.P., Z.H., W.R., Z.-G.Y., and C.-W.N. conducted the ferroelectric measurements. G.D., Z.Z., S.L., and M.L. wrote the first draft of the manuscript. All authors contributed to the discussion of the results and the revision of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/475/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S23
Movies S1 to S5
References (*40*–*51*)

13 July 2019; accepted 30 September 2019
10.1126/science.aay7221

BIOGEOGRAPHY

Global distribution of earthworm diversity

Helen R. P. Phillips[1,2]*, Carlos A. Guerra[1,3], Marie L. C. Bartz[4], Maria J. I. Briones[5], George Brown[6], Thomas W. Crowther[7], Olga Ferlian[1,2], Konstantin B. Gongalsky[8,9], Johan van den Hoogen[7], Julia Krebs[1,2], Alberto Orgiazzi[10], Devin Routh[7], Benjamin Schwarz[11], Elizabeth M. Bach[12,13], Joanne Bennett[1,3], Ulrich Brose[1,14], Thibaud Decaëns[15], Birgitta König-Ries[1,16], Michel Loreau[17], Jérôme Mathieu[18], Christian Mulder[19], Wim H. van der Putten[20,21], Kelly S. Ramirez[20], Matthias C. Rillig[22,23], David Russell[24], Michiel Rutgers[25], Madhav P. Thakur[20], Franciska T. de Vries[26], Diana H. Wall[12,13], David A. Wardle[27], Miwa Arai[28], Fredrick O. Ayuke[29], Geoff H. Baker[30], Robin Beauséjour[31], José C. Bedano[32], Klaus Birkhofer[33], Eric Blanchart[34], Bernd Blossey[35], Thomas Bolger[36,37], Robert L. Bradley[31], Mac A. Callaham[38], Yvan Capowiez[39], Mark E. Caulfield[40], Amy Choi[41], Felicity V. Crotty[42,43], Andrea Dávalos[35,44], Darío J. Diaz Cosin[45], Anahí Dominguez[32], Andrés Esteban Duhour[46], Nick van Eekeren[47], Christoph Emmerling[48], Liliana B. Falco[49], Rosa Fernández[50], Steven J. Fonte[51], Carlos Fragoso[52], André L. C. Franco[12], Martine Fugère[31], Abegail T. Fusilero[53,54], Shaieste Gholami[55], Michael J. Gundale[56], Mónica Gutiérrez López[45], Davorka K. Hackenberger[57], Luis M. Hernández[58], Takuo Hishi[59], Andrew R. Holdsworth[60], Martin Holmstrup[61], Kristine N. Hopfensperger[62], Esperanza Huerta Lwanga[63,64], Veikko Huhta[65], Tunsisa T. Hurisso[51,66], Basil V. Iannone III[67], Madalina Iordache[68], Monika Joschko[69], Nobuhiro Kaneko[70], Radoslava Kanianska[71], Aidan M. Keith[72], Courtland A. Kelly[51], Maria L. Kernecker[73], Jonatan Klaminder[74], Armand W. Koné[75], Yahya Kooch[76], Sanna T. Kukkonen[77], H. Lalthanzara[78], Daniel R. Lammel[23,79], Iurii M. Lebedev[8,9], Yiqing Li[80], Juan B. Jesus Lidon[45], Noa K. Lincoln[81], Scott R. Loss[82], Raphael Marichal[83], Radim Matula[84], Jan Hendrik Moos[85,86], Gerardo Moreno[87], Alejandro Morón-Ríos[88], Bart Muys[89], Johan Neirynck[90], Lindsey Norgrove[91], Marta Novo[45], Visa Nuutinen[92], Victoria Nuzzo[93], Mujeeb Rahman P[94], Johan Pansu[95,96], Shishir Paudel[82], Guénola Pérès[97], Lorenzo Pérez-Camacho[98], Raúl Piñeiro[99], Jean-François Ponge[100], Muhammad Imtiaz Rashid[101,102], Salvador Rebollo[98], Javier Rodeiro-Iglesias[103], Miguel Á. Rodríguez[104], Alexander M. Roth[105,106], Guillaume X. Rousseau[58,107], Anna Rozen[108], Ehsan Sayad[55], Loes van Schaik[109], Bryant C. Scharenbroch[110], Michael Schirrmann[111], Olaf Schmidt[37,112], Boris Schröder[22,113], Julia Seeber[114,115], Maxim P. Shashkov[116,117], Jaswinder Singh[118], Sandy M. Smith[119], Michael Steinwandter[115], José A. Talavera[120], Dolores Trigo[45], Jiro Tsukamoto[121], Anne W. de Valença[122], Steven J. Vanek[51], Iñigo Virto[123], Adrian A. Wackett[124], Matthew W. Warren[125], Nathaniel H. Wehr[126], Joann K. Whalen[127], Michael B. Wironen[128], Volkmar Wolters[129], Irina V. Zenkova[130], Weixin Zhang[131], Erin K. Cameron[132,133]†, Nico Eisenhauer[1,2]†

Soil organisms, including earthworms, are a key component of terrestrial ecosystems. However, little is known about their diversity, their distribution, and the threats affecting them. We compiled a global dataset of sampled earthworm communities from 6928 sites in 57 countries as a basis for predicting patterns in earthworm diversity, abundance, and biomass. We found that local species richness and abundance typically peaked at higher latitudes, displaying patterns opposite to those observed in aboveground organisms. However, high species dissimilarity across tropical locations may cause diversity across the entirety of the tropics to be higher than elsewhere. Climate variables were found to be more important in shaping earthworm communities than soil properties or habitat cover. These findings suggest that climate change may have serious implications for earthworm communities and for the functions they provide.

Soils harbor high biodiversity and are responsible for a wide range of ecosystem functions and services upon which terrestrial life depends (*1*). Despite calls for large-scale biogeographic studies of soil organisms (*2*), global biodiversity patterns remain relatively unknown, with most efforts focused on soil microbes (*3*–*5*). Consequently, the drivers of soil biodiversity, particularly soil fauna, remain unknown at the global scale.

Furthermore, our ecological understanding of global biodiversity patterns [e.g., latitudinal diversity gradients (*6*)] is largely based on the distribution of aboveground taxa. Yet many soil organisms have shown global diversity patterns that differ from aboveground organisms (*3*, *7*–*9*), although the patterns often depend on the size of the soil organism (*10*).

Here, we analyzed global patterns in earthworm diversity, total abundance, and total biomass (hereafter "community metrics"). Earthworms are considered ecosystem engineers (*11*) in many habitats and also provide a variety of vital ecosystem functions and services (*12*). The provisioning of ecosystem functions by earthworms likely depends on the abundance, biomass, and ecological group of the earthworm species (*13*, *14*). Consequently, understanding global patterns in community metrics for earthworms is critical for predicting how changes in their communities may alter ecosystem functioning.

Small-scale field studies have shown that soil properties such as pH and soil carbon influence earthworm diversity (*11*, *15*, *16*). For example, lower pH values constrain the diversity of earthworms by reducing calcium availability (*17*), and soil carbon provides resources that sustain earthworm diversity and population sizes (*11*). Alongside many interacting soil properties (*15*), a variety of other drivers can shape earthworm diversity, such as climate and habitat cover (*11*, *18*, *19*). However, to date, no framework has integrated a comprehensive set of environmental drivers of earthworm communities to identify the most important ones at a global scale.

Previous reviews suggested that earthworms may have high diversity across the tropics as a result of high endemism (*10*). However, this high regional diversity may not be captured by local-scale metrics. Alternatively, in the temperate region, local diversity may be higher (*20*) but may include fewer endemic species (*10*). We anticipate that earthworm community metrics (particularly diversity) will not follow global patterns seen aboveground, and instead, as seen across Europe (*15*), will increase with latitude. This finding would be consistent with previous studies at regional scales, which showed that the species richness of earthworms increases with latitude (*19*). Because of the relationship among earthworm communities, habitat cover, and soil properties on local scales, we expect soil properties (e.g., pH and soil organic carbon) to be key environmental drivers of earthworm communities.

Here, we present global maps predicting local diversity (number of species), abundance, and biomass. (We use "local" in the sense of site-level: a location of one or more samples that adequately captured the earthworm community.) We collated 180 datasets from the literature and unpublished field studies (164 and 16, respectively) to create a dataset spanning 57 countries (all continents except Antarctica) and 6928 sites (Fig. 1A). We explored spatial patterns of earthworm communities and determined the environmental drivers that shape earthworm biodiversity. We then used the relationships between earthworm community metrics and environmental drivers (table S1) to predict local earthworm communities across the globe.

Three generalized linear mixed-effects models were constructed, one for each of the three community metrics: species richness (calculated within a site), abundance per m^2, and biomass per m^2. Each model contained 12 environmental variables as main effects (table S2), which were grouped into six themes; "soil," "precipitation," "temperature," "water retention," "habitat cover," and "elevation" [habitat cover and some soil variables were measured in the field; the remaining variables were extracted from global data layers

based on the geographic coordinates of the sites (*14*)]. Within each theme, each model contained interactions between the variables. After model simplification, all models retained most of the original variables, but some interactions were removed (table S3).

Consistent with previous results (*20*), local earthworm diversity predictions based on global environmental data layers resulted in estimates of one to four species per site across most of the terrestrial surface (Fig. 1B) (mean, 2.42 species; SD, 2.19). Most of the boreal and subarctic regions were predicted to have low values of species richness, which is in line with aboveground biodiversity patterns (*21*, *22*). However, low local diversity also occurred in subtropical and tropical areas, such as Brazil, India, and Indonesia, in contrast to commonly

[1]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany. [2]Institute of Biology, Leipzig University, 04103 Leipzig, Germany. [3]Institute of Biology, Martin Luther University Halle-Wittenberg, 06108 Halle (Saale), Germany. [4]Universidade Positivo, Curitiba, PR 81280-330, Brazil. [5]Departamento de Ecología y Biología Animal, Universidad de Vigo, 36310 Vigo, Spain. [6]Embrapa Forestry, Colombo, PR 83411-000, Brazil. [7]Crowther Lab, Department of Environmental Systems Science, Institute of Integrative Biology, ETH Zürich, 8092 Zürich, Switzerland. [8]A. N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow 119071, Russia. [9]M. V. Lomonosov Moscow State University, Moscow 119991, Russia. [10]European Commission, Joint Research Centre, Ispra, Italy. [11]Biometry and Environmental System Analysis, University of Freiburg, 79106 Freiburg, Germany. [12]Department of Biology, Colorado State University, Fort Collins, CO 80523, USA. [13]Global Soil Biodiversity Initiative and School of Global Environmental Sustainability, Colorado State University, Fort Collins, CO 80523, USA. [14]Institute of Biodiversity, Friedrich Schiller University Jena, 07743 Jena, Germany. [15]CEFE, UMR 5175, CNRS–Univ Montpellier–Univ Paul–Valéry–EPHE–SupAgro Montpellier–INRA–IRD, 34293 Montpellier Cedex 5, France. [16]Institute of Computer Science, Friedrich Schiller University Jena, 07743 Jena, Germany. [17]Centre for Biodiversity Theory and Modeling, Theoretical and Experimental Ecology Station, CNRS, 09200 Moulis, France. [18]Sorbonne Université, CNRS, UPEC, Paris 7, INRA, IRD, Institut d'Ecologie et des Sciences de l'Environnement de Paris, F-75005 Paris, France. [19]Department of Biological, Geological and Environmental Sciences, University of Catania, 95124 Catania, Italy. [20]Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), 6700 AB Wageningen, Netherlands. [21]Laboratory of Nematology, Department of Plant Sciences, Wageningen University and Research, 6708 PB Wageningen, Netherlands. [22]Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), 14195 Berlin, Germany. [23]Institute of Biology, Freie Universität Berlin, 14195 Berlin, Germany. [24]Department of Soil Zoology, Senckenberg Museum for Natural History Görlitz, 02826 Görlitz, Germany. [25]National Institute for Public Health and the Environment, Bilthoven, Netherlands. [26]Institute of Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1012 WX Amsterdam, Netherlands. [27]Asian School of the Environment, Nanyang Technological University, 639798 Singapore. [28]Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization, Tsukuba 305-8604, Japan. [29]Department of Land Resource Management and Agricultural Technology (LARMAT), College of Agriculture and Veterinary Sciences, University of Nairobi, Nairobi 00625, Kenya. [30]CSIRO Health and Biosecurity, Canberra, ACT 2601, Australia. [31]Département de Biologie, Université de Sherbrooke, Sherbrooke J1K 2R1, Canada. [32]Geology Department, FCEFQyN, ICBIA-CONICET (National Scientific and Technical Research Council), National University of Río Cuarto, X5804 BYA Río Cuarto, Argentina. [33]Department of Ecology, Brandenburg University of Technology, 03046 Cottbus, Germany. [34]Eco&Sols, University of Montpellier, IRD, CIRAD, INRA, Montpellier SupAgro, 34060 Montpellier, France. [35]Department of Natural Resources, Cornell University, Ithaca, NY 14853, USA. [36]School of Biology and Environmental Science, University College Dublin, Belfield, Dublin 4, Ireland. [37]UCD Earth Institute, University College Dublin, Belfield, Dublin 4, Ireland. [38]USDA Forest Service, Southern Research Station, Athens, GA 30602, USA. [39]UMR 1114 "EMMAH," INRA, Site Agroparc, 84914 Avignon, France. [40]Farming Systems Ecology, Wageningen University and Research, 6700 AK Wageningen, Netherlands. [41]Faculty of Forestry, University of Toronto, Toronto, ON M5S 3B3, Canada. [42]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3EE, UK. [43]School of Agriculture, Food and Environment, Royal Agricultural University, Cirencester GL7 6JS, UK. [44]Department of Biological Sciences, SUNY Cortland, Cortland, NY 13045, USA. [45]Biodiversity, Ecology and Evolution, Faculty of Biology, Complutense University of Madrid, 28040 Madrid, Spain. [46]Laboratorio de Ecología, Instituto de Ecología y Desarrollo Sustentable, Universidad Nacional de Luján, 6700 Luján, Argentina. [47]Louis Bolk Institute, 3981 AJ Bunnik, Netherlands. [48]Department of Soil Science, Faculty of Regional and Environmental Sciences, University of Trier, Campus II, 54286 Trier, Germany. [49]Ciencias Básicas, Instituto de Ecología y Desarrollo Sustentable–INEDES, Universidad Nacional de Luján, 6700 Luján, Argentina. [50]Institute of Evolutionary Biology (CSIC–Universitat Pompeu Fabra), 08003 Barcelona, Spain. [51]Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80523, USA. [52]Biodiversity and Systematic Network, Instituto de Ecología A.C., Xalapa 91070, Mexico. [53]Department of Biological Science and Environmental Studies, University of the Philippines–Mindanao, Barangay Mintal, 8000 Davao City, Philippines. [54]Laboratory of Environmental Toxicology and Aquatic Ecology, Environmental Toxicology Unit (GhEnToxLab), Ghent University (UGent), Campus Coupure, Ghent, Belgium. [55]Razi University, Kermanshah, Iran. [56]Forest Ecology and Management, Swedish University of Agricultural Sciences, 90183 Umeå, Sweden. [57]Department of Biology, J. J. Strossmayer University of Osijek, 31000 Osijek, Croatia. [58]Agricultural Engineering, Postgraduate Program in Agroecology, Maranhão State University, 65055-310 São Luís, Brazil. [59]Faculty of Agriculture, Kyushu University, 949 Ohkawauchi, Shiiba 883-0402, Japan. [60]Minnesota Pollution Control Agency, St. Paul, MN 55155, USA. [61]Department of Bioscience, Aarhus University, 8600 Silkeborg, Denmark. [62]Biological Sciences, Northern Kentucky University, Highland Heights, KY 41099, USA. [63]Agricultura Sociedad y Ambiente, Colegio de la Frontera Sur, Ciudad Industrial, Lerma, Campeche 24500, Mexico. [64]Soil Physics and Land Management Degradation, Wageningen University and Research, 6708 PB Wageningen, Netherlands. [65]Department of Biological and Environmental Science, University of Jyväskylä, 40014 Jyväskylä, Finland. [66]College of Agriculture, Environmental and Human Sciences, Lincoln University of Missouri, Jefferson City, MO 65101, USA. [67]School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611, USA. [68]Sustainable Development and Environment Engineering, Banat's University of Agricultural Sciences and Veterinary Medicine "King Michael the 1st of Romania," 300645 Timisoara, Romania. [69]Experimental Infrastructure Platform, Leibniz Centre for Agricultural Landscape Research (ZALF), 15374 Müncheberg, Germany. [70]Faculty of Food and Agricultural Sciences, Fukushima University, Kanayagawa 1, Fukushima City, Japan. [71]Department of Environmental Management, Faculty of Natural Sciences, Matej Bel University, Banská Bystrica, Slovakia. [72]Centre for Ecology and Hydrology, Bailrigg, Lancaster LA1 4AP, UK. [73]Land Use and Governance, Leibniz Centre for Agricultural Landscape Research (ZALF), 15374 Müncheberg, Germany. [74]Department of Ecology and Environmental Science, Climate Impacts Research Centre, Umeå University, 90187 Umeå, Sweden. [75]UR Gestion Durable des Sols, UFR Sciences de la Nature, Université Nangui Abrogoua, 02 BP 801 Abidjan 02, Côte d'Ivoire. [76]Faculty of Natural Resources and Marine Sciences, Tarbiat Modares University, 46417-76489, Noor, Mazandaran, Iran. [77]Production Systems, Horticulture Technologies, Natural Resources Institute Finland, 40500 Jyväskylä, Finland. [78]Department of Zoology, Pachhunga University College, Aizawl 796001, India. [79]Soil Science, ESALQ-USP, Universidade de São Paulo, Piracicaba 13418, Brazil. [80]College of Agriculture, Forestry and Natural Resource Management, University of Hawai'i, Hilo, HI 96720, USA. [81]Tropical Plant and Soil Sciences, College of Tropical Agriculture and Human Resources, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA. [82]Department of Natural Resource Ecology and Management, Oklahoma State University, Stillwater, OK 74078, USA. [83]UR Systèmes de pérennes, CIRAD, Univ Montpellier, 34398 Montpellier, France. [84]Department of Forest Ecology, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, 165 21 Prague, Czech Republic. [85]Department of Soil and Environment, Forest Research Institute of Baden-Wuerttemberg, 79100 Freiburg, Germany. [86]Thuenen-Institute of Organic Farming, 23847 Westerau, Germany. [87]Forestry School–INDEHESA, University of Extremadura, 10600 Plasencia, Spain. [88]Conservación de la Biodiversidad, El Colegio de la Frontera Sur, 24500 Campeche, Mexico. [89]Department of Earth and Environmental Sciences, KU Leuven, 3001 Leuven, Belgium. [90]Research Institute for Nature and Forest, 9500 Geraardsbergen, Belgium. [91]School of Agricultural, Forest and Food Sciences, Bern University of Applied Sciences, 3052 Zollikofen, Switzerland. [92]Soil Ecosystems, Natural Resources Institute Finland (Luke), 31600 Jokioinen, Finland. [93]Natural Area Consultants, 1 West Hill School Road, Richford, NY 13835, USA. [94]Department of Zoology, Pocker Sahib Memorial Orphanage College, Tirurangadi, Malappuram, Kerala, India. [95]CSIRO Ocean and Atmosphere, Lucas Heights, NSW 2234, Australia. [96]UMR7144 Adaptation et Diversité en Milieu Marin, Station Biologique de Roscoff, CNRS-Sorbonne Université, 29688 Roscoff, France. [97]UMR SAS, INRA, Agrocampus Ouest, 35000 Rennes, France. [98]Ecology and Forest Restoration Group, Department of Life Sciences, University of Alcalá, 28801 Alcalá De Henares, Spain. [99]Computing, ESEI, Vigo, Edf. Politécnico–Campus As Lagoas, 32004 Ourense, Spain. [100]Adaptations du Vivant, CNRS UMR 7179, Muséum National d'Histoire Naturelle, 91800 Brunoy, France. [101]Centre of Excellence in Environmental Studies, King Abdulaziz University, Jeddah 21589, Saudi Arabia. [102]Environmental Sciences, COMSATS University Islamabad, Sub-campus Vehari, Vehari 61100, Pakistan. [103]Departamento de Informática, Escuela Superior de Ingeniería Informática, Universidad de Vigo, 36310 Vigo, Spain. [104]Group of Global Change Ecology and Evolution (GloCEE), Department of Life Sciences, University of Alcalá, 28805 Alcalá de Henares, Spain. [105]Department of Forest Resources, University of Minnesota, St. Paul, MN 55101, USA. [106]Friends of the Mississippi River, 101 East Fifth Street, St. Paul, MN 55108, USA. [107]Postgraduate Program in Biodiversity and Conservation, Federal University of Maranhão, 65080-805 São Luís, Brazil. [108]Institute of Environmental Sciences, Jagiellonian University, 30-087 Kraków, Poland. [109]Institute of Ecology, Technical University of Berlin, 10587 Berlin, Germany. [110]College of Natural Resources, University of Wisconsin, Stevens Point, WI 54481, USA. [111]Engineering for Crop Production, Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), 14469 Potsdam, Germany. [112]UCD School of Agriculture and Food Science, University College Dublin, Belfield, Ireland. [113]Landscape Ecology and Environmental Systems Analysis, Institute of Geoecology, Technische Universität Braunschweig, 38106 Braunschweig, Germany. [114]Department of Ecology, University of Innsbruck, 6020 Innsbruck, Austria. [115]Institute for Alpine Environment, Eurac Research, 39100 Bozen/Bolzano, Italy. [116]Laboratory of Ecosystem Modeling, Institute of Physicochemical and Biological Problems in Soil Sciences, Russian Academy of Science, Pushchino 142290, Russia. [117]Laboratory of Computational Ecology, Institute of Mathematical Problems of Biology–Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino 142290, Russia. [118]Post Graduate Department of Zoology, Khalsa College Amritsar, Amritsar 143002, India. [119]John H. Daniels Faculty of Architecture, Landscape and Design, University of Toronto, Toronto, ON M5S 3B3, Canada. [120]Department of Animal Biology, University of La Laguna, 38200 La Laguna, Spain. [121]Faculty of Agriculture, Kochi University, Monobe Otsu 200, Nankoku 783-8502, Japan. [122]Food and Agriculture, WWF-Netherlands, 3708 JB Zeist, Netherlands. [123]Dpto. Ciencias, IS-FOOD, Universidad Pública de Navarra, Edificio Olivos–Campus Arrosadia, 31006 Pamplona, Spain. [124]Soil, Water and Climate, University of Minnesota, St. Paul, MN 55108, USA. [125]Earth Innovation Institute, 98 Battery Street, San Francisco, CA 94111, USA. [126]Department of Natural Resources and Environmental Management, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA. [127]Natural Resource Sciences, McGill University, Ste-Anne-de-Bellevue H9X 3V9, Canada. [128]The Nature Conservancy, Arlington, VA 22203, USA. [129]Department of Animal Ecology, Justus Liebig University, 35392 Giessen, Germany. [130]Laboratory of Terrestrial Ecosystems, Kola Science Centre, Institute of the North Industrial Ecology Problems, Apatity 184211, Russia. [131]Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions (Henan University), Ministry of Education, College of Environment and Planning, Henan University, Kaifeng 475004, China. [132]Department of Environmental Science, Saint Mary's University, Halifax, Nova Scotia, Canada. [133]Faculty of Biological and Environmental Sciences, University of Helsinki, FI 00014 Helsinki, Finland.
***Corresponding author. Email: helen.phillips@idiv.de** †These authors contributed equally to this work.

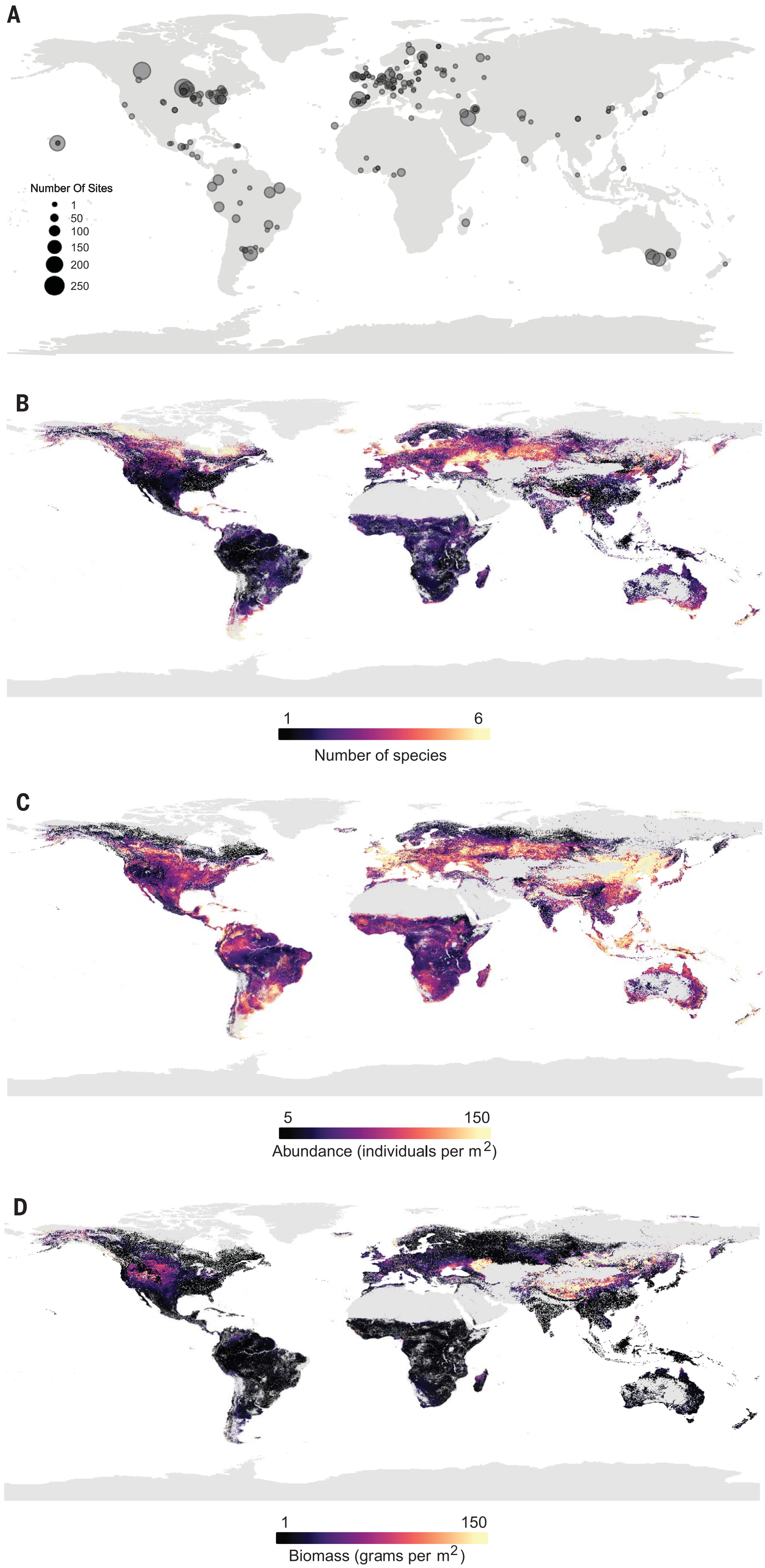


Fig. 1. Global distribution of earthworm diversity. (**A**) Black dots represent the center of a "study" used in at least one of the three models (species richness, total abundance, and total biomass). The size of the dot corresponds to the number of sites within the study. Opaqueness is for visualization purposes only. (**B** to **D**) The globally predicted values of (B) species richness (within site), (C) total abundance, and (D) total biomass. Yellow indicates high diversity; dark purple, low diversity. Gray areas are habitat cover categories that lacked samples.

observed aboveground patterns, such as the latitudinal gradient in plant diversity (*22*). This pattern could be due to different relationships with climate variables. For example, although plant diversity increases with potential evapotranspiration (PET) (*22*), earthworm diversity tended to decrease with increasing PET (table S3). In addition, soil properties, which are typically not included in models of aboveground diversity, can play a role in determining earthworm communities (*11*, *15*, *23*). For instance, litter availability and soil nutrient content are important regulators of earthworm diversity, with oligotrophic forest soils having more epigeic species and eutrophic soils more endogeics (*23*). Furthermore, tropical regions with higher decomposition rates have fewer soil organic resources and lower local earthworm diversity (Fig. 1B and table S3), dominated by endogeic species, which have specific digestion systems that allow them to feed on low-quality soil organic matter (*11*, *14*, *20*).

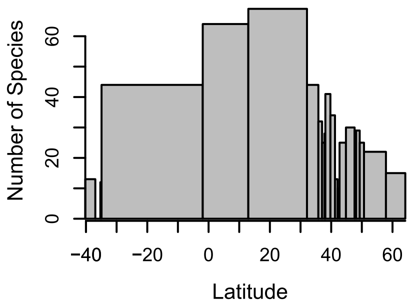


Fig. 2. The number of unique species within each latitudinal zone, when the number of sites within each zone is comparable. The width of the bar shows the latitude range of the sites/zones.

High local species richness was found at mid-latitudes, such as the southern tip of South America, the southern regions of Australia and New Zealand, Europe (particularly north of the Black Sea), and the northeastern United States. Although this pattern contrasts with latitudinal diversity patterns found in many aboveground organisms (*6*, *24*), it is consistent with patterns found in some belowground organisms [ectomycorrhizal fungi (*3*), bacteria (*5*)], but not all [arbuscular mycorrhizal fungi (*25*), oribatid mites (*26*)]. Such mismatches between above- and belowground biodiversity have been predicted (*1*, *7*) but not shown across the globe for soil fauna at the local scale.

The patterns seen here could in part be a result of glaciation in the last ice age, as well as human activities. Temperate regions (mid- to high latitudes) that were previously glaciated were likely recolonized by earthworm species with high dispersal capabilities and large geographic ranges (*19*) and through human-mediated dispersal ["anthropochorous" earthworms (*16*)]. Thus, temperate communities could have high local diversity, as seen here, but those species would be widely distributed, resulting in lower regional diversity relative to local diversity. In the tropics, which did not experience glaciation, the opposite may be true. Specific locations may have individual species that are highly endemic, but these species are not widely distributed (table S4). This high local endemism would result in low local diversity (as found here) and high regional diversity [as suggested by (*10*)] relative to that low local diversity. When the numbers of unique species within latitudinal zones that had equal numbers of sites were calculated (i.e., a regional richness that accounted for sampling effort), there appeared to be a regional latitudinal diversity gradient (Fig. 2). Even with a sampling bias (table S4), regional richness in the tropics was greater than in the temperate regions, despite low local diversity. These results should be interpreted with caution, given the latitude span of the tropical zones. However, the underlying data suggest that endemism of earthworms and β-diversity within the tropics (*27*) may be considerably higher than within the well-sampled temperate region (table S4). Therefore, it is likely that the tropics harbor more species overall.

The predicted total abundance of the local community of earthworms typically ranged between 5 and 150 individuals per m^2 across the globe, in line with other estimates (*28*) (Fig. 1C; mean, 77.89 individuals per m^2; SD, 98.94). There was a slight tendency for areas of higher total abundance to be in temperate areas, such as Europe (particularly the UK, France, and Italy), New Zealand, and part of the Pampas and surrounding region (South America), rather than the tropics. Lower total abundance occurred in many of the tropical and subtropical regions, such as Brazil, central Africa, and parts of India. Given the positive relationship between total abundance and ecosystem function (*29*), in regions with lower earthworm abundance, such functions may be reduced or carried out by other soil taxa (*1*).

The predicted total biomass of the local earthworm community (adults and juveniles) across the globe showed extreme values (>2 kg) in 0.3% of pixels, but biomass typically ranged (97% of pixels) between 1 g and 150 g per m^2 [Fig. 1D; median, 6.69; mean, 2772.8; SD, 1,312,782; see (*14*) for additional discussion of extreme values]. The areas of high total biomass were concentrated in the Eurasian Steppe and some regions of North America. The majority of the globe showed low total biomass. In northern North America, where there are no native earthworms (*13*), high density and, in some regions, higher biomass of earthworms likely reflect the earthworm invasion of these regions. The small invasive European earthworm species encounter an enormous unused resource pool, which leads to high population sizes (*30*). On the basis of previous suggestions (*28*), we expected that

Table 1. Model validation results. Cells in boldface show the "best" value when comparing between the main models (a mixture of sampled soil properties and SoilGrids data) and models containing only SoilGrids data. Values shown are mean square error [MSE; calculated for all predicted data ("Total") and for tertiles ("Low," "Mid," "High")] following 10-fold cross-validation of the main models and models containing only SoilGrids data, as well as R^2 of the main models and SoilGrids-only models.

	Total	Low	Mid	High
MSE: Main models				
Species richness	**1.376**	0.917	0.812	**3.561**
Abundance	**17977.42**	**1720.75**	**2521.25**	**48751.51**
Biomass	3220.29	264.56	**441.25**	8783.77
MSE: SoilGrids models				
Species richness	1.385	**0.887**	**0.793**	3.716
Abundance	18775.81	1735.11	2516.13	51156.76
Biomass	**3068.00**	**199.91**	461.88	**8380.81**

	Marginal	Conditional
R^2: Main models		
Species richness	0.132	**0.748**
Abundance	0.176	**0.626**
Biomass	0.201	0.612
R^2: SoilGrids models		
Species richness	**0.142**	0.745
Abundance	**0.234**	0.643
Biomass	**0.242**	**0.650**

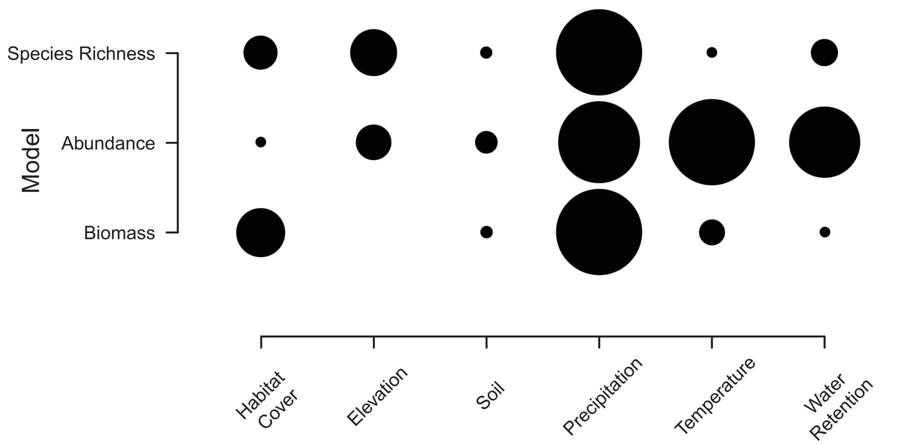


Fig. 3. The importance of the six variable themes from the three biodiversity models. Rows show the results of each model (top, species richness; middle, abundance; bottom, biomass). Columns represent the variable themes that are present in the simplified biodiversity model. The most important variable group has the largest circle. Within each row, the circle size of the other variable themes is proportional to the relative change in importance. The circle size should only be compared within a row.

earthworms would decrease in body size toward the poles, showing low biomass relative to the total abundance in temperate or boreal regions. In contrast, in tropical regions (e.g., Brazil and Indonesia) that are dominated by giant earthworms that normally occur at low densities and low species richness (*31*), we expected high biomass but low abundance. However, these patterns were not found. This could be due to the relatively small number of sample points for the biomass model (n = 3296) compared to the diversity (n = 5416) and total abundance models (n = 6358), reducing the predictive ability of the model (fig. S1C), most notably in large regions of Asia and in areas of Africa, particularly the boundaries of the Sahara Desert and the southern regions (which coincides with sites where samples are lacking). Additionally, the difficulty in consistently capturing such large earthworms in every sample may increase data variability, reducing the ability of the model to predict.

Overall, the three community metric models performed well in cross-validation (figs. S3 and S4) with relatively high R^2 values [Table 1; see (*14*) for further details and caveats]. But given the nature of such analyses, models and maps should only be used to explore broad patterns in earthworm communities and not at the fine scale, especially in relation to conservation practices (*32*).

For all three community metric models, climatic variables were the most important drivers (the "precipitation" theme being the most important for both species richness and total biomass models, and "temperature" for the total abundance model; Fig. 3). The importance of climatic variables in shaping diversity and distribution patterns at large scales is consistent with many aboveground taxa [e.g., plants (*22*), reptiles, amphibians, and mammals (*32*)] and belowground taxa [bacteria and fungi (*3*, *5*), nematodes (*33*)]. This suggests that climate-related methods and data, which are typically used by macroecologists to estimate aboveground biodiversity, may also be suitable for estimating earthworm communities. However, the strong link between climatic variables and earthworm community metrics is cause for concern, as climate will continue to change due to anthropogenic activities over the coming decades (*34*). Our findings further highlight that changes in temperature and precipitation are likely to influence earthworm diversity (*35*) and distributions (*15*), with implications for the functions that they provide (*12*). Shifts in distributions may be particularly problematic in the case of invasive earthworms, such as in areas of North America, where they can considerably change the ecosystem (*13*). However, a change in climate will most likely affect abundance and biomass of the earthworm communities before it affects diversity, as shifts in the latter depend on dispersal capabilities, which are relatively low in earthworms.

We expected that soil properties would be the most important driver of earthworm communities, but this was not the case (Fig. 3), likely because of the scale of the study. First, the importance of drivers could change at different spatial scales. Climate is driving patterns at global scales, but within climatic regions (or at the local scale), other variables may become more important (*36*). Thus, one or more soil properties may be the most important drivers of earthworm communities within each of the primary studies, rather than across them all. Second, for soil properties, the mismatch in scale between community metrics and the soil properties taken from global layers [for sites where sampled soil properties were missing (*14*)] potentially reduced the apparent importance of the theme. Habitat cover influenced the earthworm community (fig. S5, A and B), especially the composition of the three ecological groups (epigeics, endogeics, and anecics) (fig. S6) (*14*). Across larger scales, climate influences both habitat cover and soil properties, all of which affect earthworm communities. Being able to account for this indirect effect with appropriate methods and data may alter the perceived importance of soil properties and habitat cover [e.g., with pathway analysis (*37*) and standardized data]. However, our habitat cover variable did not directly consider local management (such as land use or intensity).

Our findings suggest that climate change might have substantial effects on earthworm communities and the functioning of ecosystems; any climate change–induced alteration in earthworm communities is likely to have cascading effects on other species in these ecosystems (*13*, *28*). Despite earthworm communities being controlled by environmental drivers similar to those that affect aboveground communities (*22*, *37*), these relationships result in different patterns of diversity. We highlight the need to integrate belowground organisms into biodiversity research, despite differences in the scale of sampling, if we are to fully understand large-scale patterns of biodiversity and their underlying drivers (*7*, *8*, *38*), especially if processes underlying macroecological patterns differ between aboveground and belowground diversity (*38*). The inclusion of soil taxa may alter the distribution of biodiversity hotspots and conservation priorities. For example, protected areas (*7*) may not be protecting earthworms (*7*), despite their importance as ecosystem function providers (*12*) and soil ecosystem engineers for other organisms (*11*). By modeling both realms, aboveground/belowground comparisons are possible, potentially allowing a clearer view of the biodiversity distribution of whole ecosystems.

REFERENCES AND NOTES

1. R. D. Bardgett, W. H. van der Putten, *Nature* **515**, 505–511 (2014).
2. N. Eisenhauer *et al.*, *Pedobiologia* **63**, 1–7 (2017).
3. L. Tedersoo *et al.*, *Science* **346**, 1256688 (2014).
4. M. Delgado-Baquerizo *et al.*, *Science* **359**, 320–325 (2018).
5. M. Bahram *et al.*, *Nature* **560**, 233–237 (2018).
6. H. Hillebrand, *Am. Nat.* **163**, 192–211 (2004).
7. E. K. Cameron *et al.*, *Conserv. Biol.* **33**, 1187–1192 (2019).
8. N. Fierer, M. S. Strickland, D. Liptzin, M. A. Bradford, C. C. Cleveland, *Ecol. Lett.* **12**, 1238–1249 (2009).
9. J. van den Hoogen *et al.*, *Nature* **572**, 194–198 (2019).
10. T. Decaëns, *Glob. Ecol. Biogeogr.* **19**, 287–302 (2010).
11. C. A. Edwards, Ed., *Earthworm Ecology* (CRC Press, ed. 2, 2004).
12. M. Blouin *et al.*, *Eur. J. Soil Sci.* **64**, 161–182 (2013).
13. D. Craven *et al.*, *Glob. Change Biol.* **23**, 1065–1074 (2017).
14. See supplementary materials.

15. M. Rutgers *et al*., *Appl. Soil Ecol.* **97**, 98–111 (2016).
16. P. F. Hendrix, P. J. Bohlen, *Bioscience* **52**, 801–811 (2002).
17. T. G. Piearce, *J. Anim. Ecol.* **41**, 167 (1972).
18. D. J. Spurgeon, A. M. Keith, O. Schmidt, D. R. Lammertsma, J. H. Faber, *BMC Ecol.* **13**, 46 (2013).
19. J. Mathieu, T. J. Davies, *J. Biogeogr.* **41**, 1204–1214 (2014).
20. P. Lavelle, C. Lattaud, D. Trigo, I. Barois, *Plant Soil* **170**, 23–33 (1995).
21. R. R. Dunn *et al*., *Ecol. Lett.* **12**, 324–333 (2009).
22. H. Kreft, W. Jetz, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5925–5930 (2007).
23. C. Fragoso, P. Lavelle, *Soil Biol. Biochem.* **24**, 1397–1408 (1992).
24. K. J. Gaston, T. M. Blackburn, *Pattern and Process in Macroecology* (Blackwell, 2000).
25. J. Davison *et al*., *Science* **349**, 970–973 (2015).
26. M. Maraun, H. Schatz, S. Scheu, *Ecography* **30**, 209–216 (2007).
27. T. Decaëns *et al*., *Soil Biol. Biochem.* **92**, 171–183 (2016).
28. D. C. Coleman, D. A. Crossley, P. F. Hendrix, *Fundamentals of Soil Ecology* (Elsevier, ed. 2, 2004), pp. 271–298.
29. J. W. Spaak *et al*., *Ecol. Lett.* **20**, 1315–1324 (2017).
30. N. Eisenhauer, J. Schlaghamerský, P. B. Reich, L. E. Frelich, *Biol. Invasions* **13**, 2191–2196 (2011).
31. M. A. Drumond *et al*., *Braz. J. Biol.* **73**, 699–708 (2013).
32. L. Santini *et al*., *Glob. Ecol. Biogeogr.* **27**, 968–979 (2018).
33. D. Song *et al*., *Appl. Soil Ecol.* **114**, 161–169 (2017).
34. Intergovernmental Panel on Climate Change, *Climate Change 2014 Synthesis Report Summary for Policymakers* (2014); www.ipcc.ch/site/assets/uploads/2018/02/AR5_SYR_FINAL_SPM.pdf.
35. D. K. Hackenberger, B. K. Hackenberger, *Eur. J. Soil Biol.* **61**, 27–34 (2014).
36. M. A. Bradford *et al*., *Nat. Ecol. Evol.* **1**, 1836–1845 (2017).
37. A. Rice *et al*., *Nat. Ecol. Evol.* **3**, 265–273 (2019).
38. A. Shade *et al*., *Trends Ecol. Evol.* **33**, 731–744 (2018).

ACKNOWLEDGMENTS

We thank all the reviewers who provided thoughtful and constructive feedback on this manuscript. We thank M. Winter and the sDiv team for their help in organizing the sWorm workshops, and the Biodiversity Informatics Unit (BDU) at iDiv for their assistance in making the data open access. In addition, the data providers thank all supervisors, students, collaborators, technicians, data analysts, land owners/managers, and anyone else involved with the collection, processing, and/or publication of the primary datasets. **Funding:** This work was developed during and following two sWorm workshops. H.R.P.P. and the sWorm workshops were supported by the sDiv [Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (DFG FZT 118)]. H.R.P.P., O.F., and N.E. acknowledge funding by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 677232 to N.E.). K.S.R. and W.H.v.d.P. were supported by ERC-ADV grant 323020 to W.H.v.d.P. Also supported by iDiv (DFG FZT118) Flexpool proposal 34600850 (C.A.G. and N.E.); the Academy of Finland (285882) and the Natural Sciences and Engineering Research Council of Canada (postdoctoral fellowship and RGPIN-2019-05758) (E.K.C.); DOB Ecology (T.W.C., J.v.d.H., and D.R.); ERC-AdG 694368 (M.R.); and the TULIP Laboratory of Excellence (ANR-10-LABX-41) (M.L.). In addition, data collection was funded by the Russian Foundation for Basic Research (12-04-01538-a, 12-04-01734-a, 14-44-03666-r_center_a, 15-29-02724-ofi_m, 16-04-01878-a 19-05-00245); Tarbiat Modares University; Aurora Organic Dairy; UGC(NERO) (F. 1-6/Acctt./NERO/2007-08/1485); Natural Sciences and Engineering Research Council (RGPIN-2017-05391); Slovak Research and Development Agency (APVV-0098-12); Science for Global Development through Wageningen University; Norman Borlaug LEAP Programme and International Atomic Energy Agency (IAEA); São Paulo Research Foundation - FAPESP (12/22510-8); Oklahoma Agricultural Experiment Station; INIA - Spanish Agency (SUM 2006-00012-00-0); Royal Canadian Geographical Society; Environmental Protection Agency (Ireland) (2005-S-LS-8); University of Hawai'i at Mānoa (HAW01127H; HAW01123M); European Union FP7 (FunDivEurope, 265171); U.S. Department of the Navy, Commander Pacific Fleet (W9126G-13-2-0047); Science and Engineering Research Board (SB/SO/AS-030/2013) Department of Science and Technology, New Delhi, India; Strategic Environmental Research and Development Program (SERDP) of the U.S. Department of Defense (RC-1542); Maranhão State Research Foundation (FAPEMA); Coordination for the Improvement of Higher Education Personnel (CAPES); Ministry of Education, Youth and Sports of the Czech Republic (LTT17033); Colorado Wheat Research Foundation; Zone Atelier Alpes, French National Research Agency (ANR-11-BSV7-020-01, ANR-09-STRA-02-01, ANR 06 BIODIV 009-01); Austrian Science Fund (P16027, T441); Landwirtschaftliche Rentenbank Frankfurt am Main; Welsh Government and the European Agricultural Fund for Rural Development (Project Ref. A AAB 62 03 qA731606); SÉPAQ; Ministry of Agriculture and Forestry of Finland; Science Foundation Ireland (EEB0061); University of Toronto (Faculty of Forestry); National Science and Engineering Research Council of Canada; Haliburton Forest and Wildlife Reserve; NKU College of Arts and Sciences Grant; Österreichische Forschungsförderungsgesellschaft (837393 and 837426); Mountain Agriculture Research Unit of the University of Innsbruck; Higher Education Commission of Pakistan; Kerala Forest Research Institute, Peechi, Kerala; UNEP/GEF/TSBF-CIAT Project on Conservation and Sustainable Management of Belowground Biodiversity; Ministry of Agriculture and Forestry of Finland; Complutense University of Madrid/European Union FP7 project BioBio (FPU UCM 613520); GRDC; AWI; LWRRDC; DRDC; CONICET (National Scientific and Technical Research Council) and FONCyT (National Agency of Scientific and Technological Promotion) (PICT, PAE, PIP), Universidad Nacional de Luján y FONCyT [PICT 2293 (2006)], Fonds de recherche sur la nature et les technologies du Québec (131894), Deutsche Forschungsgemeinschaft [SCHR1000/3-1, SCHR1000/6-1, 6-2 (FOR 1598), WO 670/7-1, WO 670/7-2, and SCHA 1719/1-2], CONACYT (FONDOS MIXTOS TABASCO/PROYECTO11316); NSF (DGE-0549245, DGE-0549245, DEB-BE-0909452, NSF1241932); Institute for Environmental Science and Policy at the University of Illinois at Chicago; Dean's Scholar Program at UIC; Garden Club of America Zone VI Fellowship in Urban Forestry from the Casey Tree Endowment Fund; J. E. Weaver Competitive Grant from the Nebraska Chapter of The Nature Conservancy; the College of Liberal Arts and Sciences at DePaul University; Elmore Hadley Award for Research in Ecology and Evolution from the UIC Dept. of Biological Sciences; Spanish CICYT (AMB96-1161; REN2000-0783/GLO; REN2003-05553/GLO; REN2003-03989/GLO; CGL2007-60661/BOS); Yokohama National University; MEXT KAKENHI (25220104); Japan Society for the Promotion of Science KAKENHI (25281053, 17KT0074, 25252026); ADEME (0775C0035); Ministry of Science, Innovation and Universities of Spain (CGL2017-86926-P); Syngenta Philippines; UPSTREAM; LTSER (Val Mazia/Matschertal); Marie Sklodowska Curie Postdoctoral Fellowship (747607); National Science and Technology Base Resource Survey Project of China (2018FY100306); McKnight Foundation (14-168); Program of Fundamental Researches of Presidium of Russian Academy of Sciences (AAAA-A18-118021490070-5); Brazilian National Council of Research CNPq; and French Ministry of Foreign and European Affairs. **Author contributions:** H.R.P.P. led the analysis, data curation, and writing of the original manuscript draft. C.A.G. assisted in analyses and writing of the original manuscript draft. E.K.C. and N.E. revised subsequent manuscript drafts. J.v.d.H., D.R., and T.W.C. provided additional analyses. E.K.C., N.E., and M.P.T. acquired funding for the project. J.K., K.B.G., B.S., M.L.C.B., M.J.I.B., and G.B. contributed to data curation. H.R.P.P., C.A.G., M.L.C.B., M.J.I.B., G.B., O.F., A.O., E.M.B., J.B., U.B., T.D., F.T.d.V., B.K.-R., M.L., J.M., C.M., W.H.v.d.P., K.S.R., M.C.R., D.R., M.R., M.P.T., D.H.W., D.A.W., E.K.C., and N.E. contributed to the project conceptualization. All authors reviewed and edited the final draft manuscript. The majority of the authors provided data for the analyses. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data and analysis code are available on the iDiv Data repository (DOI: 10.25829/idiv.1804-5-2593) and GitHub (https://github.com/helenphillips/GlobalEWDiversity; DOI: 10.5281/zenodo.3386456).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/480/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1 to S4
References (*39*–*76*)

1 April 2019; accepted 10 September 2019
10.1126/science.aax4851

NANOMATERIALS

High-surface-area corundum by mechanochemically induced phase transformation of boehmite

Amol P. Amrute[1]*, Zbigniew Łodziana[2], Hannah Schreyer[1], Claudia Weidenthaler[1], Ferdi Schüth[1]*

In its nanoparticulate form, corundum (α-Al_2O_3) could lead to several applications. However, its production into nanoparticles (NPs) is greatly hampered by the high activation energy barrier for its formation from cubic close-packed oxides and the sporadic nature of its nucleation. We report a simple synthesis of nanometer-sized α-Al_2O_3 (particle diameter ~13 nm, surface areas ~140 $m^2\ g^{-1}$) by the mechanochemical dehydration of boehmite (γ-AlOOH) at room temperature. This transformation is accompanied by severe microstructural rearrangements and might involve the formation of rare mineral phases, diaspore and tohdite, as intermediates. Thermodynamic calculations indicate that this transformation is driven by the shift in stability from boehmite to α-Al_2O_3 caused by milling impacts on the surface energy. Structural water in boehmite plays a crucial role in generating and stabilizing α-Al_2O_3 NPs.

Corundum (α-Al_2O_3) in high-surface-area nanoparticle (NP) form (*1*, *2*) would enable a number of applications. For example, in auto-exhaust catalysis, gamma alumina (γ-Al_2O_3) is used in wash-coat formulations, but α-Al_2O_3 has higher mechanical stability and could be used directly as support without wash coating if available as NPs (*3*, *4*). It could also provide improved hydrothermal stability when water is present in the exhaust gas. For ammonia synthesis, the presence of α-Al_2O_3 in the catalyst formulation was shown to increase catalyst activity by threefold (*5*). Corundum is widely used in ceramic applications, including dental implants, prostheses, and high-speed cutting tools, where high fracture strength and toughness is desired (*6*, *7*). The use of nanoparticulate, fine-grained corundum precursors would allow the manufacture of ceramics with improved fracture toughness and high density at reduced sintering temperatures, enabling energy savings (*6*, *7*).

A number of studies—mostly theoretical (*1*, *8*–*10*) but a few experimental (*1*, *2*)—have addressed the thermodynamics and synthesis of corundum NPs. However, none of the efforts made has thus far resulted in a method for the fabrication of high-purity corundum with surface areas reaching or even exceeding 100 $m^2\ g^{-1}$ (*1*, *2*, *11*). One reason for this lack of success may be the high activation energy barrier (485 kJ mol^{-1}), requiring temperatures above 1473 K, to facilitate a rearrangement of oxide ions from the cubic close-packed structure found in transition aluminas to the hexagonally close-packed form in corundum (Fig. 1, middle) (*9*, *12*). This transition typically leads to substantial mass transfer during the transformation that is associated with a loss in surface area, generally to values well below 10 $m^2\ g^{-1}$.

Thermodynamic calculations coupled to adsorption microcalorimetry and thermogravimetric analyses by Navrotsky et al. suggest that at room temperature, γ-Al_2O_3 could actually be the thermodynamically most stable phase for NPs, corresponding to surface areas of 100 to 200 $m^2\ g^{-1}$ (*1*, *13*). At 800 K, γ-Al_2O_3 might become stable even at lower surface areas. Thus, a high-temperature process, which tends to lead to the formation of the most stable phase, should never lead to high-surface-area α-Al_2O_3. However, density functional theory simulations by Nørskov et al. showed that the surface energy, and thus the stability, of an alumina polymorph depends strongly on the size of the crystallite and degree of hydroxylation (*9*).

Accordingly, some hydroxylated surfaces of alumina polymorphs, especially θ-Al_2O_3, but also α-Al_2O_3, should have negative surface energies, and the nanocrystalline form of alumina should become thermodynamically most stable. In fact, a later study by Navrotsky et al. also showed lower surface energy for the hydroxylated form of α-Al_2O_3 than for a similarly hydroxylated γ-Al_2O_3 irrespective of the surface area (*13*). Hence, α-Al_2O_3 with a high surface area >100 $m^2\ g^{-1}$ could be accessible and synthesizable, but numerous studies to date have proven unsuccessful at producing it through the thermal route at 1500 K.

The topotactic transformation of diaspore (α-AlOOH) to α-Al_2O_3 produces corundum at moderate temperatures of 800 to 873 K (Fig. 1, top) (*1*, *14*), but the synthesis of stable diaspore requires very harsh conditions (723 K, 1200 bar, 35 days) (*15*, *16*). Mechanochemical reactions such as that induced by ball milling, which finds increasing attention as a simple, economical, and scalable method for the preparation of nanomaterials and for carrying out chemical reactions (*17*–*20*), have also been reported as a method to obtain α-Al_2O_3 from γ-Al_2O_3 (*21*, *22*). Nonequilibrium processes involved in breaking and making chemical bonds by mechanical energy input enable this phase transformation to take place at room temperature. However, γ-Al_2O_3 undergoes continuous agglomeration until it fully transforms to α-Al_2O_3, which severely reduces the specific surface area (by a factor of >2.2) (*21*).

The observed effect might be related to the so-called "milling equilibrium" between cold welding, leading to the formation of larger agglomerates, and fragmentation, producing smaller particles (*23*). The dominance of one process over the other determines the eventual size of the produced particles. As γ-Al_2O_3 is softer (bulk modulus, B_{VR} = 196.9 GPa, and shear modulus, G_{VR} = 97.45 GPa; see table S1), it could undergo excessive cold welding, which would cease with the formation of harder α-Al_2O_3 (B_{VR} = 239.6 GPa and G_{VR} = 138.6 GPa), as harder materials are difficult to cold weld and also difficult to further fragmentize without external aids.

We first approached this problem by performing ball milling of γ-Al_2O_3 in the presence of process control agents (PCAs), which are known to limit cold welding and enhance fragmentation (*23*). Stearic acid, alcohols, and heavier alkanes are commonly used as PCAs (*23*). However, we opted for water as an environmentally benign and cost-effective PCA. The ball-milling experiments were performed in a vibration mill. Typically, 1 g of precursor powder is charged in the tungsten carbide (WC) vessel with three WC balls (see section S2 of the supplementary materials for more details). The effect of abrasion of milling balls, the vessel, or both caused contamination of the samples with WC after milling. WC was selectively removed by an oxidative leaching step for all the samples obtained using the WC milling jar and media (supplementary text and fig. S1). Ball milling of γ-Al_2O_3 (denoted as γ-Al_2O_3-1, S_{BET} = 109 $m^2\ g^{-1}$; see fig. S2A for morphology) in the presence of 1 to 5 weight % (wt %) H_2O for 120 min fully transformed the material into α-Al_2O_3, but the surface area dropped to 70 $m^2\ g^{-1}$ (fig. S2B). This decrease in surface area is similar to its ball milling without water (fig. S2C). A further increase in H_2O to 15 wt % resulted in no phase transformation of γ-Al_2O_3, suggesting that an excess amount of water in the system is detrimental because it might reduce the energy transfer during the milling impact.

These results suggest that high-surface-area γ-Al_2O_3 is a suitable precursor for the production of α-Al_2O_3 with moderately high surface area, but that alternative precursors might be better. Ball milling was also reported to convert boehmite to α-Al_2O_3 (*24*), but there is no information on the surface area, intermediates, or agglomeration phenomena, and thus it remains unclear whether boehmite could lead to

[1]Department of Heterogeneous Catalysis, Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany. [2]INP, Polish Academy of Sciences, ul. Radzikowskiego 152, PL- 31-342 Kraków, Poland.
*Corresponding author. Email: amrute@kofo.mpg.de (A.P.A.); schueth@kofo.mpg.de (F.S.)

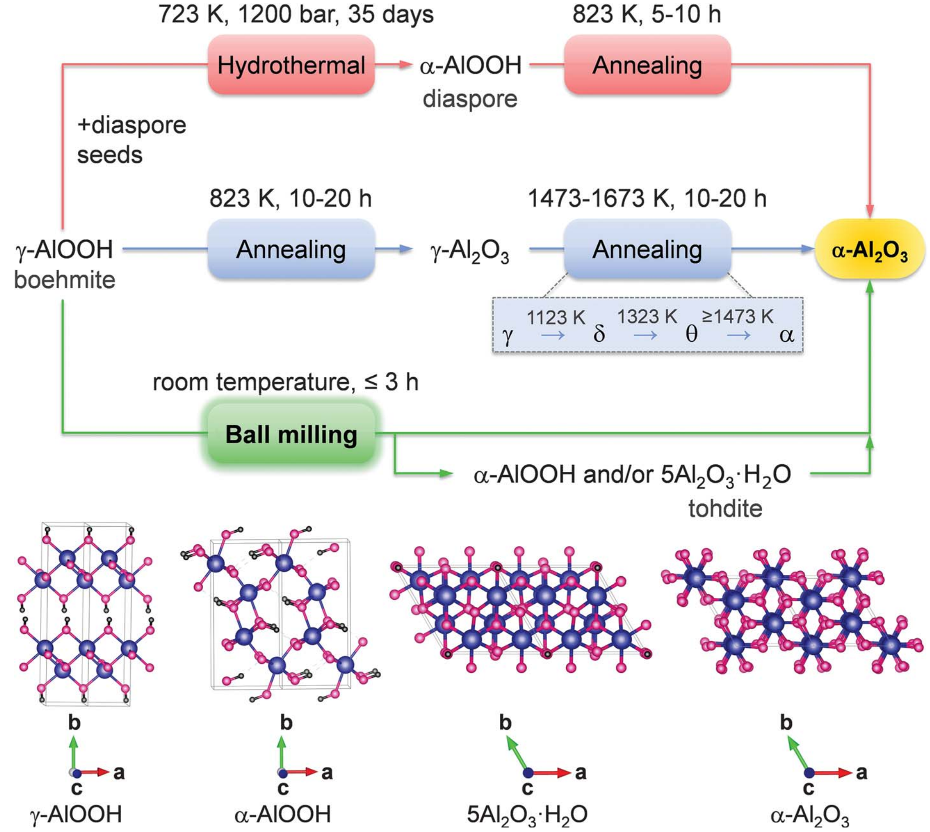


Fig. 1. Methods for the synthesis of α-Al_2O_3. In contrast to the harsh conditions needed in conventional routes (top and middle), ball milling (bottom) enables us to synthesize α-Al_2O_3 in the nanoparticulate form at room temperature from easily accessible boehmite precursors. Bottom insets show the crystal structures (two-unit cells) of boehmite [crystal system: orthorhombic; space group: *Cmcm* (*27*)], diaspore [crystal system: orthorhombic; space group: *Pbnm* (*28*)], tohdite [crystal system: hexagonal; space group: $P6_3mc$ (*10*)], and α-Al_2O_3 [crystal system: hexagonal-trigonal; space group: $R\bar{3}c$ no. 167 (*29*)] along the c-axis. In boehmite, aluminum and oxygen atoms form double layers of octahedra that sandwich hydrogen atoms in a zigzag manner. The diaspore structure consists of the "double rutile strings" formed by the aluminum-centered oxygen octahedra and involves hexagonal close packing of the oxygen ions (AB-AB…). The tohdite structure is composed of oxygen planes stacked in the sequence of A-B-A-C-A… and a unit cell contains 10 Al atoms, eight of which occupy octahedral and two tetrahedral sites. In the α-Al_2O_3 structure, each Al^{3+} center is octahedrally coordinated and the oxygen ions form a hexagonal close-packed structure with the aluminum ions filling two-thirds of the octahedral interstices. Atom colors are as follows: Al, blue; O, pink; and H, black.

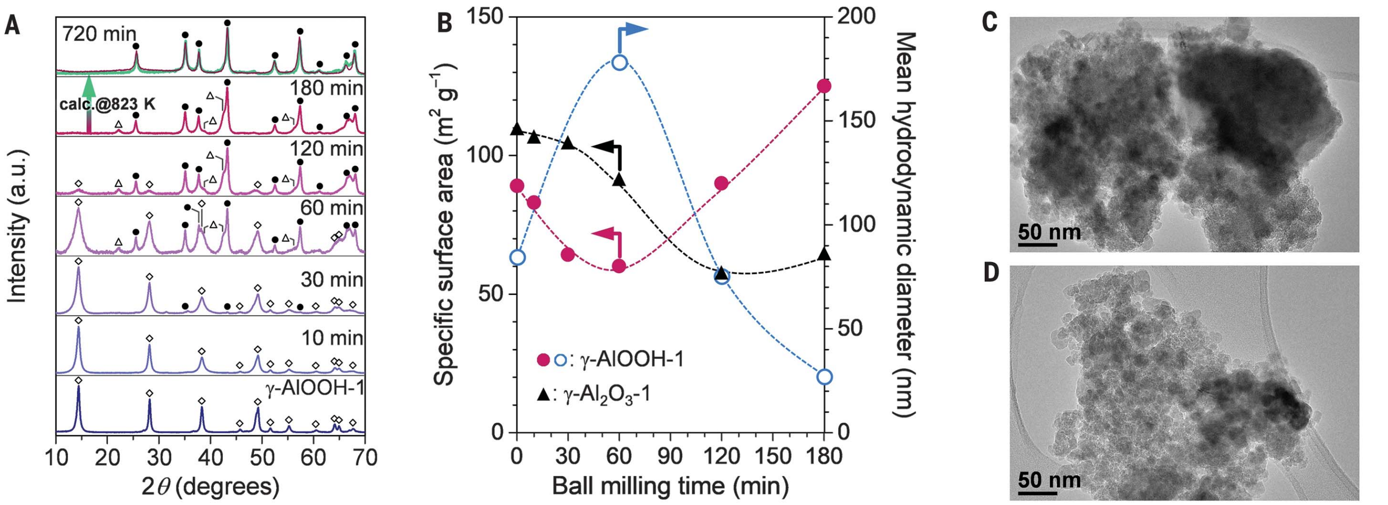


Fig. 2. Evolution of the α-Al_2O_3 phase and changes in the surface area upon ball milling of boehmite. (**A**) Diffractograms of γ-AlOOH-1 before and after ball milling at room temperature in the vibration mill. Milling conditions: WC jar (25 cm^3), 3 × 12 mm WC balls, 1 g of powder charge, f_{Mill} = 25 Hz, t = 10 to 720 min. Identified crystalline phases: ◇, γ-AlOOH; ●, α-Al_2O_3; and △, α-AlOOH. (**B**) Surface area and hydrodynamic diameter of particles versus milling time for γ-AlOOH (circles) and γ-Al_2O_3 (triangles) precursors. (**C** and **D**) TEM images of 60-min ball-milled γ-AlOOH-1 (**C**) and 180-min ball-milled γ-AlOOH-1 (D). The ball milling of boehmite leads to the formation of α-Al_2O_3 with diaspore and/or tohdite (fig. S4A) intermediates. The formation of the latter relies on the milling conditions. The phase transformation to α-Al_2O_3 proceeds through severe microstructural rearrangements.

high-surface-area corundum. Boehmite attracted our interest also because it formally has one water molecule in its structure per alumina (i.e., 15 wt % water with respect to alumina) and can be economically produced, for example, by a modified Bayer process of bauxite refining (*25*). With this material, an effective milling impact can still be achieved, and the structural water of boehmite, which will be released upon its dehydration to alumina, could act as a PCA to limit cold welding or to facilitate the fragmentation of agglomerates.

On the basis of these considerations, we ball milled a boehmite sample (denoted as γ-AlOOH-1; S_{BET} = 89 $m^2\ g^{-1}$; see fig. S3 for morphology) for different times using a WC jar and media (Fig. 2A). The boehmite phase remained mostly unaltered upon ball milling for up to 30 min,

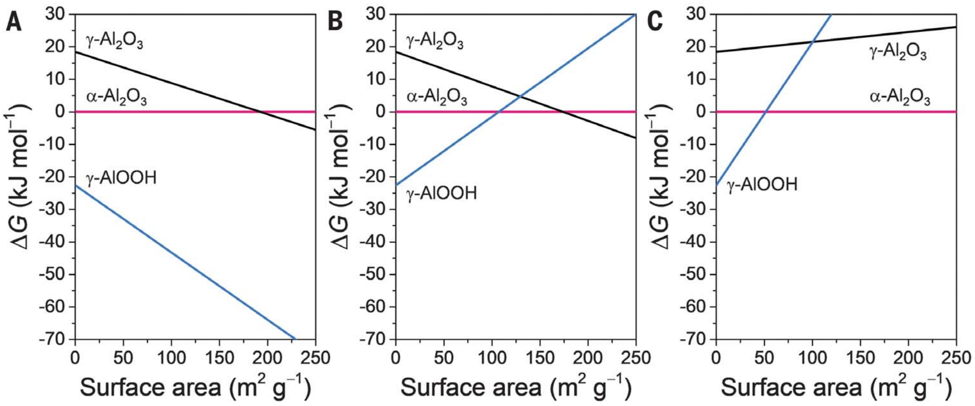


Fig. 3. Free energy of γ-Al_2O_3 and γ-AlOOH with respect to α-Al_2O_3 as a function of surface area at standard state (T = 298.15 K and P = 1 bar). Data in (**A**) correspond to the conditions discussed in (*1*), where at high surface area, γ-Al_2O_3 is thermodynamically more stable than α-Al_2O_3. (**B**) Modification of boehmite surface in the milling process effectively increases the surface energy of this compound (above 4 J m^{-2}; fig. S7), which drives the conversion to α-Al_2O_3 above 108 m^2 g^{-1}. (**C**) Alumina surface hydroxylation changes the mutual stability of the phases, where α-Al_2O_3 becomes the most stable phase above 52 m^2 g^{-1}. For the corresponding data for diaspore and tohdite, see figs. S8 and S9.

Table 1. Surface area, phase composition, and domain size of boehmite before and after ball milling using a WC jar and media in the vibration mill and after calcination of the ball-milled sample.

Entry	Sample	Before milling			After 180 min of milling*		After calcination of the milled sample at 823 K		
		S_{BET} (m^2 g^{-1})	Phase†	Domain size (nm)‡	S_{BET} (m^2 g^{-1})	Phase	S_{BET} (m^2 g^{-1})	Phase	Domain size (nm)
1	γ-AlOOH-1	89	b	17	125	c + d	120	c	18
2	γ-AlOOH-2	64	b	20	115	c + d	136	c	18
3	γ-AlOOH-3	10	b	52	103	c + d	130	c	19
4	γ-AlOOH-4, dried	366	b	4	132	c + d	140	c	18

*Other ball-milling conditions can be found in the legend to Fig. 2A. †Identified crystalline phases: b, boehmite; d, diaspore; and c, corundum. ‡Domain size determined by the application of the Scherrer equation to the corresponding XRD pattern (see section S3 in the supplementary materials for more details).

but x-ray diffraction (XRD) reflections corresponding to α-Al_2O_3 and α-AlOOH appeared after 60 min of milling. Further continuation of milling for 180 min fully converted boehmite to a mixture of corundum (major, ~70 to 80%) and diaspore (minor, ~20 to 30%). Similar results were obtained upon ball milling of boehmite with a steel milling jar and media (fig. S4A), except instead of diaspore, another mineral phase, tohdite ($5Al_2O_3 \cdot H_2O$, which is structurally similar to α-Al_2O_3, see Fig. 1, bottom insets, and rearranges to the latter upon dehydration), was detected.

The observation of these rare phases by ball milling of boehmite was surprising because their formation normally requires extremely harsh conditions (*16*, *26*). However, in view of the thermodynamic analysis (see below), the formation of such phases is possible for the high-surface-area systems, which was confirmed by performing experiments under variable conditions (fig. S5). In the ball milling of γ-Al_2O_3, no intermediate phase was observed (*21*, *22*), which we confirmed by milling γ-Al_2O_3-1 for different times (fig. S2C).

Ball milling of boehmite for 720 min led to the formation of pure α-Al_2O_3 (Fig. 2A and fig. S4A), showing that these intermediates can also be transformed to the desired corundum phase (alternatively, the remaining diaspore after 180 min of milling can be topotactically converted to α-Al_2O_3 by a mild calcination at 823 K; Fig. 2A, green diffractogram). This result indicates that the conversion of these phases is slower compared with that of boehmite, likely because they are mechanically harder (table S1). Increased milling frequency (30 Hz) led to a faster transformation of boehmite and intermediates to α-Al_2O_3 (fig. S4B and table S2).

These results show that the formation of intermediates depends on the milling conditions and that high impact energy favors the nucleation of the α-Al_2O_3 phase. This finding was substantiated by ball milling of boehmite in the planetary mill using steel, sintered corundum, or zirconia milling jars and media. The higher impact and friction forces involved in planetary milling led to the direct formation of α-Al_2O_3 (see fig. S6). Notably, milling equipment made of corundum can also be used to exclude the presence of any foreign material by abrasion.

We obtained deeper insight into the phase transformation discussed above from the thermodynamic analysis of the involved phases (see section S4 of the supplementary materials). At standard state, γ-AlOOH should be the most stable phase irrespective of the surface area (Fig. 3A and table S3). For the dense bulk phases, the thermodynamic stability decreases in the order γ-AlOOH > α-Al_2O_3 > γ-Al_2O_3 and above a surface area of 190 m^2 g^{-1}, γ-Al_2O_3 becomes more stable than corundum. This trend agrees with those reported for these alumina polymorphs (*13*). Assuming that water vacancies are the primary defects among others (see section S4 in the supplementary materials), the modification of γ-AlOOH by ball milling effectively increases its surface energy and reverses the mutual stability orders as follows: α-Al_2O_3 > γ-AlOOH > γ-Al_2O_3 (Fig. 3, B and C, and fig. S7). At a surface area >108 m^2 g^{-1} for the anhydrous state and 52 m^2 g^{-1} for the hydroxylated state, α-Al_2O_3 becomes the thermodynamically most stable phase, which provides the driving force for the boehmite-to-corundum transformation.

Mechanochemical conversion is a nonequilibrium process in which intermediates can be observed. Comparison of the surface energies of diaspore and tohdite with boehmite and corundum suggests that boehmite has the largest thermodynamic driving force for its conversion under our milling conditions (figs. S8C and S9). Thus, intermediate phases such as diaspore or tohdite could also form, together with the most stable α-Al_2O_3. Diaspore or tohdite would also ultimately convert to α-Al_2O_3, but the driving force for their conversion is lower and requires longer times or higher milling intensity.

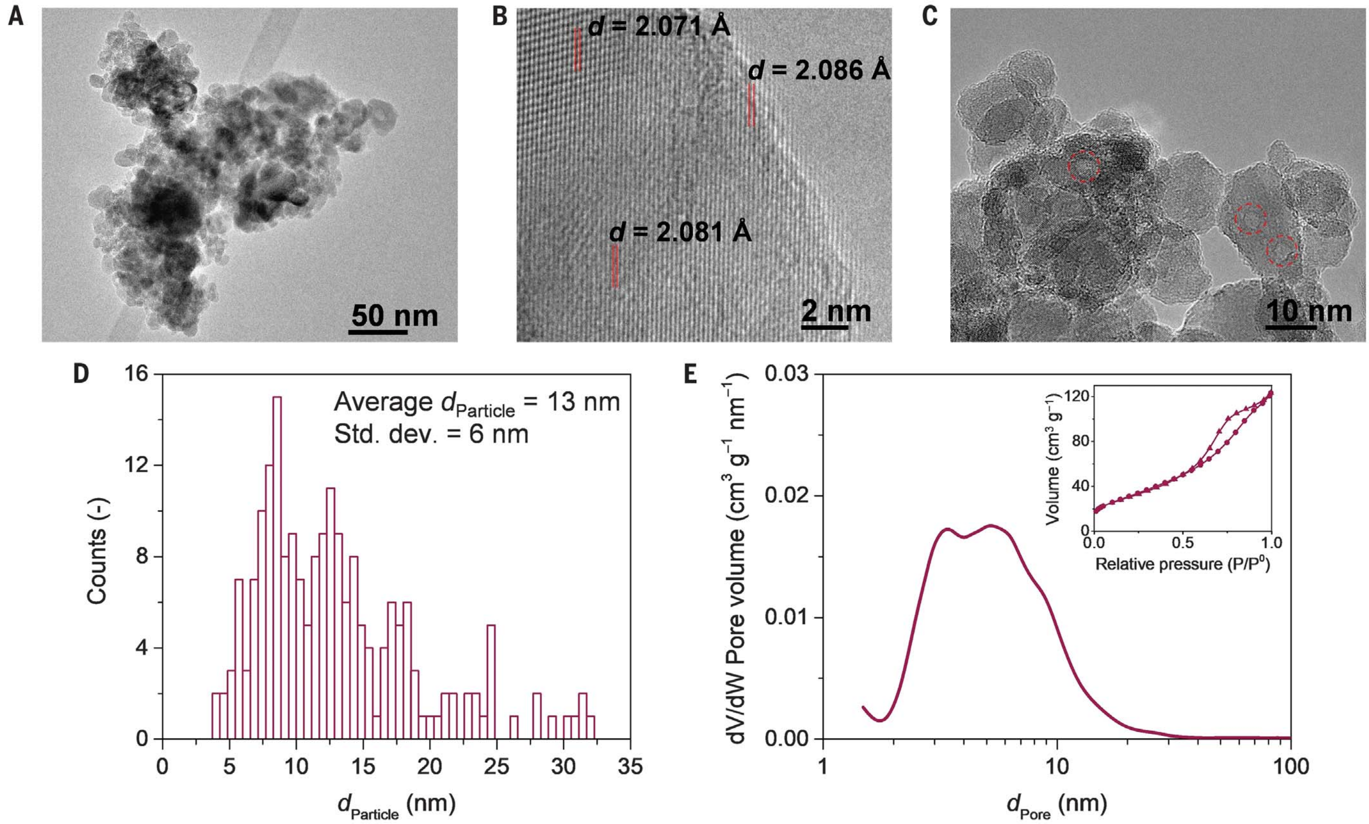


Fig. 4. Properties of nanoparticulate α-Al_2O_3 derived by ball milling of boehmite. (**A**) TEM and (**B** and **C**) high-resolution TEM of nanoparticulate α-Al_2O_3 obtained by 180-min ball milling of γ-AlOOH-1, followed by calcination at 823 K. (**D**) Particle size ($d_{Particle}$) distribution from TEM images of nanoparticulate α-Al_2O_3. (**E**) Pore size distribution determined by the application of density functional theory model to the N_2 isotherm. Inset in (E) shows N_2 adsorption-desorption isotherms. Ball-milling–derived α-Al_2O_3 from boehmite is composed of rounded NPs ($d_{Particle}$ = 13 nm) with surface roughness and defects and pores in the range of 3 to 10 nm.

To determine whether these phase changes involve agglomeration, fragmentation, or both phenomena, we further characterized the ball-milled boehmite and γ-Al_2O_3 samples. The surface area gradually decreased with milling time for γ-Al_2O_3 and reached a plateau after 120 min (full conversion to α-Al_2O_3; Fig. 2B). This result is in agreement with previous work (*21*). Likewise, boehmite also displayed a loss in surface area during the first 60 min of milling, which suggests that agglomeration was also prominent for boehmite in the initial phase of milling. A faster decrease in surface area for boehmite compared with γ-Al_2O_3 likely reflects their different mechanical properties (table S1); boehmite is softer and prone to cold weld faster than γ-Al_2O_3.

However, in contrast to γ-Al_2O_3, the surface area was recovered after 60 min for the boehmite precursor (Fig. 2B; see table S2 for steel milling jar and media). To understand this difference, the samples were characterized by dynamic light scattering (Fig. 2B). An increase in the hydrodynamic diameter during the first hour of milling and a subsequent decrease upon further milling indicated that the sample was agglomerated during the first hour of milling, and then these agglomerates were fragmented into smaller NPs. This process was visualized by transmission electron microscopy (TEM) analysis (Fig. 2, C and D). Agglomeration of the sample milled for 60 min was seen, and finely rounded NPs were detected after 180 min of milling. Thus, the structural water in boehmite seems to play a crucial role in generating and stabilizing the corundum NPs. Water released during the dehydration of boehmite to α-Al_2O_3 could create fissures in the agglomerates that facilitate their disintegration during milling, and the formed NPs of α-Al_2O_3 would be stabilized through hydroxylation of the surface by the released water. Hydroxylation has been demonstrated to lower the surface energy of α-Al_2O_3 (*13*). Further evidence highlighting the effect of structural water was derived by ball milling of boehmite at elevated temperatures (373 to 473 K) (fig. S10).

The α-Al_2O_3 NPs derived by ball milling using WC jar and balls, followed by mild calcination, were further characterized in depth (Fig. 4). TEM images showed that the NPs ranged in size from 4 to 32 nm, with an arithmetic average particle size ($d_{Particle}$) of 13 nm (standard deviation = 6 nm) (Fig. 4, A and D). The domain size determined by the application of the Scherrer equation to the corresponding XRD pattern was 18 nm (Table 1). High-resolution TEM showed that the NPs display a crystalline rather than an amorphous surface (Fig. 4B). Determination of the *d* spacing at different spots provided an average value of 2.079 Å, which agrees well with the 2.085 Å *d* spacing for the (113) plane of α-Al_2O_3.

The TEM images also showed some irregularity, surface roughness, or defects as well as some pores on the NPs (Fig. 4C). Pore size determined from N2 sorption analysis using the density functional theory model indicated that pores in the size range from 3 to 10 nm were present (Fig. 4E), which might include a contribution from interparticle pores. The surfaces were also highly dehydrated (fig. S11). Thus, the observed surface area can be attributed to nanometer-sized particles and a combination of surface roughness, defects, and pores. Similar results were obtained for other boehmite precursors of a different nature (i.e., different particle size, surface area, and water contents; see Table 1, figs. S12 to S14, and corresponding captions). In all cases, full boehmite conversion was achieved after 180 min of milling, with surface areas for α-Al_2O_3 exceeding 130 $m^2\ g^{-1}$.

Boehmite and related compounds were found to be highly effective and economically accessible

precursors to produce phase-pure nanocrystalline α-Al_2O_3 with surface areas exceeding 120 $m^2\ g^{-1}$ through mechanochemistry. These results may have high practical relevance, for instance, for the optimization of alumina processing in ceramics and for the design of stable industrial catalysts (e.g., as supports for automotive or Fischer–Tropsch synthesis catalysts).

REFERENCES AND NOTES

1. J. M. McHale, A. Auroux, A. J. Perrotta, A. Navrotsky, *Science* **277**, 788–791 (1997).
2. R. M. Laine, J. C. Marchal, H. P. Sun, X. Q. Pan, *Nat. Mater.* **5**, 710–712 (2006).
3. M. M. Bhasin, M. S. Jarrell, Catalyst comprising substantially pure alpha-alumina carrier for treating exhaust gases. U.S. Patent 5856263 (1999).
4. B. Frank, G. Emig, A. Renken, *Appl. Catal. B* **19**, 45–57 (1998).
5. B. Lin *et al.*, *ACS Catal.* **9**, 1635–1644 (2019).
6. X. Teng, H. Liu, C. Huang, *Mater. Sci. Eng. A* **452-453**, 545–551 (2007).
7. O. L. Ighodaro, O. I. Okoli, *Int. J. Appl. Ceram. Technol.* **5**, 313–323 (2008).
8. K. C. Hass, W. F. Schneider, A. Curioni, W. Andreoni, *Science* **282**, 265–268 (1998).
9. Z. Łodziana, N.-Y. Topsøe, J. K. Nørskov, *Nat. Mater.* **3**, 289–293 (2004).
10. M. Digne, P. Sautet, P. Raybaud, H. Toulhoat, E. Artacho, *J. Phys. Chem. B* **106**, 5155–5162 (2002).
11. W. L. Suchanek, J. M. Garcés, *CrystEngComm* **12**, 2996–3002 (2010).
12. C. J.-P. Steiner, D. P. H. Hasselman, R. M. Spriggs, *J. Am. Ceram. Soc.* **54**, 412–413 (1971).
13. R. H. R. Castro, S. V. Ushakov, L. Gengembre, D. Gouvea, A. Navrotsky, *Chem. Mater.* **18**, 1867–1872 (2006).
14. L. Löffler, W. Mader, *J. Am. Ceram. Soc.* **86**, 534–540 (2003).
15. H. Yanagida, G. Yamaguchi, *J. Ceram. Assoc. Jpn.* **74**, 94–100 (1966).
16. T. Tsuchida, K. Kodaira, *J. Mater. Sci.* **25**, 4423–4426 (1990).
17. P. Baláž *et al.*, *Chem. Soc. Rev.* **42**, 7571–7637 (2013).
18. A. Stolle, T. Szuppa, S. E. S. Leonhardt, B. Ondruschka, *Chem. Soc. Rev.* **40**, 2317–2329 (2011).
19. K. Ralphs, C. Hardacre, S. L. James, *Chem. Soc. Rev.* **42**, 7701–7718 (2013).
20. M. Danielis *et al.*, *Angew. Chem. Int. Ed.* **57**, 10212–10216 (2018).
21. P. A. Zieliński, R. Schulz, S. Kaliaguine, A. Van Neste, *J. Mater. Res.* **8**, 2985–2992 (1993).
22. S. R. Chauruka *et al.*, *Chem. Eng. Sci.* **134**, 774–783 (2015).
23. Y. F. Zhang, L. Lu, S. M. Yap, *J. Mater. Process. Technol.* **89-90**, 260–265 (1999).
24. A. Tonejc, A. M. Tonejc, D. Bagović, C. Kosanović, *Mater. Sci. Eng. A* **181–182**, 1227–1231 (1994).
25. D. Panias, I. Paspaliaris, *Erzmetall* **56**, 75–81 (2003).
26. M. Okumiya, G. Yamaguchi, O. Yamada, S. Ono, *Bull. Chem. Soc. Jpn.* **44**, 418–423 (1971).
27. A. N. Christensen *et al.*, *Acta Chem. Scand. A* **36a**, 303–308 (1982).
28. F. J. Ewing, *J. Chem. Phys.* **3**, 203–207 (1935).
29. I. Levin, D. Brandon, *J. Am. Ceram. Soc.* **81**, 1995–2012 (1998).

ACKNOWLEDGMENTS

The electron microscopy department at the Max-Planck-Institut für Kohlenforschung is gratefully acknowledged for the electron microscopy and the EDX analyses. We thank G. Prieto for providing boehmite samples (γ-AlOOH-1 and γ-AlOOH-4). PL-Grid is acknowledged for CPU allocation. **Funding:** Max-Planck-Institut für Kohlenforschung is acknowledged for financial support. **Author contributions:** F.S. conceived the idea of investigating ball milling to produce high-surface-area α-Al_2O_3. A.P.A. and F.S. designed a strategy to study boehmite compounds. A.P.A. synthesized the starting materials for ball milling, designed and executed ball-milling experiments, and performed characterizations. Z.Ł. formulated the thermodynamic model, performed the calculations of elastic and thermodynamic properties, and wrote the corresponding part of the text. H.S. first detected the transformation of γ-Al_2O_3 to α-Al_2O_3 in experiments to synthesize supported catalysts in the ball mill and performed some of the experiments involving ball milling of γ-Al_2O_3. C.W. assisted in the measurement of selected samples on a PANalytical X'Pert Pro diffractometer and performed Rietveld refinement of selected samples. A.P.A. and F.S. wrote the manuscript. All authors commented on the manuscript. **Competing interests:** The authors have filed a patent application on the synthesis process to obtain high-surface-area alpha-alumina. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/485/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S3
References (*30*–*40*)

14 March 2019; resubmitted 22 July 2019
Accepted 27 September 2019
10.1126/science.aaw9377

MUTATION

Higher-fitness yeast genotypes are less robust to deleterious mutations

Milo S. Johnson[1,2,3], Alena Martsul[4], Sergey Kryazhimskiy[4]*, Michael M. Desai[1,2,3,5]*

Natural selection drives populations toward higher fitness, but second-order selection for adaptability and mutational robustness can also influence evolution. In many microbial systems, diminishing-returns epistasis contributes to a tendency for more-fit genotypes to be less adaptable, but no analogous patterns for robustness are known. To understand how robustness varies across genotypes, we measure the fitness effects of hundreds of individual insertion mutations in a panel of yeast strains. We find that more-fit strains are less robust: They have distributions of fitness effects with lower mean and higher variance. These differences arise because many mutations have more strongly deleterious effects in faster-growing strains. This negative correlation between fitness and robustness implies that second-order selection for robustness will tend to conflict with first-order selection for fitness.

The dynamics and outcomes of adaptive evolution depend on the genetic variation available to a population. Because mutations interact epistatically, the availability and strength of beneficial and deleterious mutations can vary across genotypes (*1*). As "first-order" natural selection drives populations toward higher fitness, "second-order" selection can favor genotypes with better prospects for future evolution, steering populations into regions of the fitness landscape with better uphill prospects (i.e., that are more adaptable) or into "flatter" regions with fewer or less steep downhill paths (i.e., that are more robust to deleterious mutations) (*2*–*5*). Although second-order selection has been observed in microbes, viruses, and digital organisms (*4*–*8*), our understanding of how and why genotypes differ in their robustness and adaptability is incomplete.

Many laboratory microbial populations display a consistent pattern of declining adaptability, such that less-fit genotypes adapt more rapidly than more-fit genotypes (*9*–*13*). This is at least partially explained by "diminishing-returns" epistasis, in which individual beneficial mutations become less beneficial in more-fit genetic backgrounds (*12*–*15*). In contrast to adaptability, genetic variation in mutational robustness has not been systematically characterized, and no analogs to diminishing returns or the rule of declining adaptability are known. There is evidence for both antagonistic and synergistic epistasis between pairs of deleterious mutations (*16*–*19*), but little is known about how the entire distribution of fitness effects (DFE) of deleterious mutations changes across different genetic backgrounds and whether robustness, like adaptability, depends systematically on fitness.

There are several ways to define and measure mutational robustness (*2*–*4*). In this work, we consider only the single-step mutational neighborhood of a genotype. We define robustness on the basis of the DFE of single mutations and refer to strains in which these mutations are more deleterious on average as less robust (though the appropriate precise definition of mutational robustness can depend on the full DFE and the specific population genetic question). To understand genetic variation in single-step mutational robustness, we aim to measure how the DFE varies across genotypes and how these differences arise from epistasis at the level of individual mutations. To do so, we would ideally like to measure the fitness effects of identical large sets of random mutations in multiple genotypes. This would allow us to compare the DFE of these mutations across genotypes and to identify the genetic basis of differences in robustness.

To this end, we developed a pipeline to measure the effects of sets of specific insertion mutations in a panel of *Saccharomyces cerevisiae* genotypes (Fig. 1). Briefly, we transform yeast strains with transposon mutagenesis libraries derived from (*20*), in which each plasmid is tagged by multiple specific DNA barcodes. Homology-directed repair then creates the same set of transposon insertion mutations

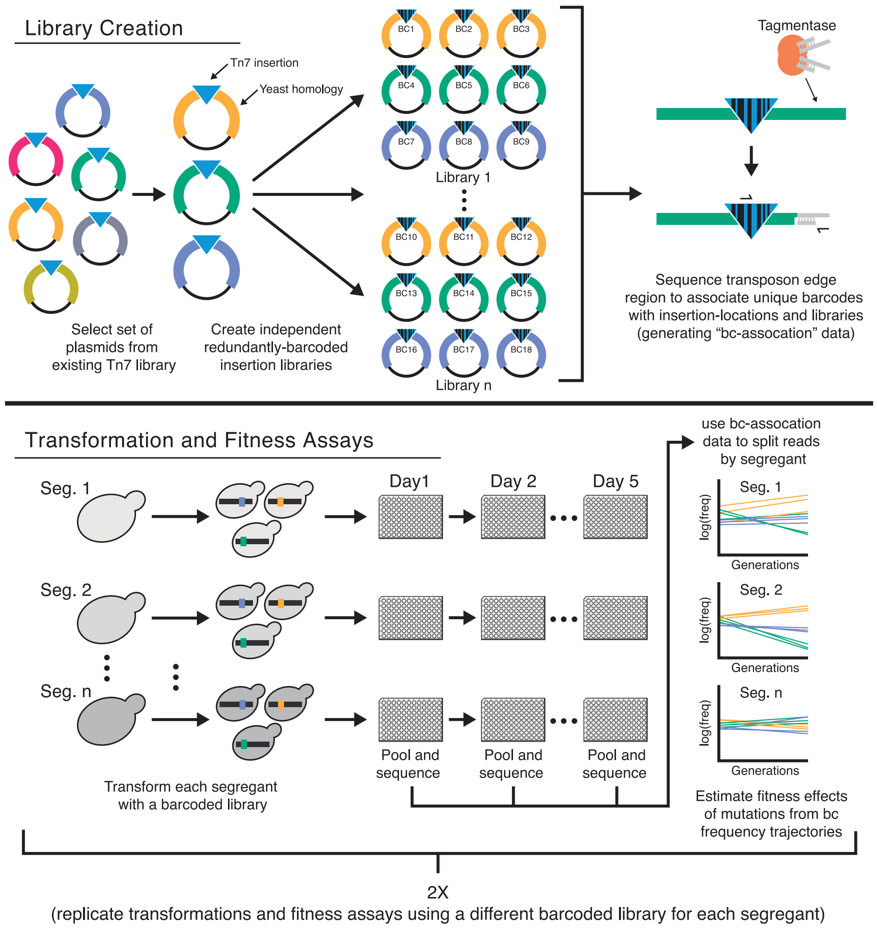


Fig. 1. Schematic of mutagenesis and fitness-assay pipeline. Plasmids with different colors indicate different regions of homology from the yeast genome. BC, barcode; Seg., segregant; Tn7, transposon 7.

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [2]Quantitative Biology Initiative, Harvard University, Cambridge, MA 02138, USA. [3]NSF-Simons Center for Mathematical and Statistical Analysis of Biology, Harvard University, Cambridge, MA 02138, USA. [4]Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA. [5]Department of Physics, Harvard University, Cambridge, MA 02138, USA.
*Corresponding author. Email: mdesai@oeb.harvard.edu (M.M.D.); skryazhi@ucsd.edu (S.K.)

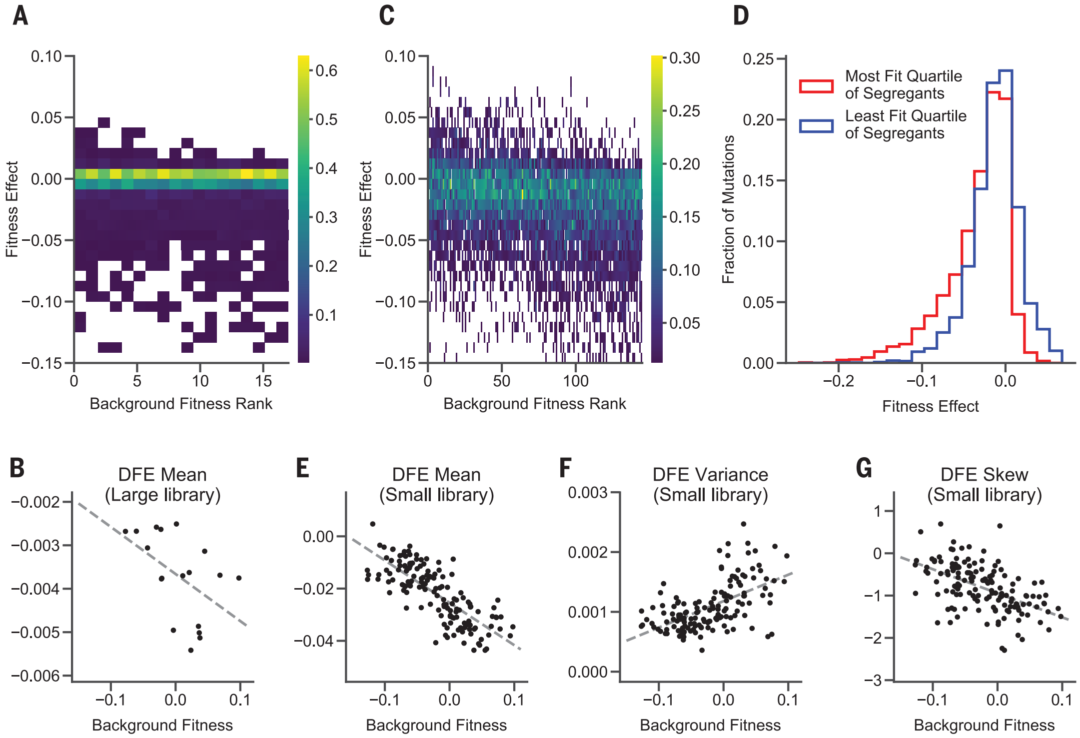


Fig. 2. DFEs of insertion mutations. (**A**) DFEs in the large library experiment. Segregants are organized by background fitness. Color represents the fraction of mutations for each segregant in each fitness-effect bin (scale bar at right). (**B**) Relationship between background fitness and the mean of the DFE for the large library experiment (P = 0.03, two-sided t test). (**C**) DFEs in the small library experiment. Color represents the fraction of mutations for each segregant in each fitness-effect bin (scale bar at right). (**D**) Combined DFE of most-fit and least-fit quartile of segregants in the small library experiment. (**E** to **G**) Relationship between background fitness and DFE statistics for the small library experiment ($P = 4.13 \times 10^{-31}$, $P = 7.83 \times 10^{-14}$, $P = 4.08 \times 10^{-12}$, respectively, two-sided t test).

in each strain. We propagate the resulting mutant pools in batch culture and measure barcode-frequency trajectories to estimate the fitness effect of each mutation in each strain (*21*) (figs. S1 and S2). Using this approach, we conducted two experiments to measure mutational robustness in F1 segregants derived from a yeast cross between a laboratory and a wine strain (BY and RM). These segregants differ at more than 35,000 loci and have previously been sequenced and phenotyped (*22*). They vary in fitness (across a 22.5% range) in our focal environment, and those with lower fitness are more adaptable (*9*).

In the first, "large library" experiment, we transformed 18 randomly chosen segregants from (*9*) with libraries consisting of 1147 mutations (fig. S3) (*21*). Although these are not a random sample of all naturally occurring mutations in yeast, they represent an unbiased set of genomic disruptions. Because of gene essentiality, differences in cloning or transformation efficiency, or other complications, we were unable to measure the fitness effect of every mutation in every segregant [analysis of the effects of missing data is in (*21*) and fig. S4]. We successfully measured the effects of 710 mutations in at least one segregant (on average, 414 per segregant). Of these 710, 457 (64%) had no detectable fitness effects in any segregant. Most remaining mutations were deleterious, and their DFE varied systematically across strains. Specifically, the mean fitness effect of these mutations decreases with the background fitness of the segregant (i.e., more-fit segregants tend to be less robust with respect to random insertion mutations, P = 0.03, two-sided t test) (Fig. 2, A and B, and fig. S5).

Because we only measured fitness effects across 18 segregants in our large library experiment, and 64% of the mutations were indistinguishable from neutral, we were unable to detect more-subtle changes in the DFE or to connect DFE-level changes to patterns of epistasis for individual mutations. To address this limitation, in our second experiment we transformed a larger group of 163 randomly chosen segregants with a "small" library, created by selecting a subset of 91 insertions from our large library that had significant fitness effects in the largest number of segregants (*21*). This small library filters out insertions with undetectable or rare effects but is otherwise unbiased.

Again, we find that the mean of the DFE decreases as segregant fitness increases (Fig. 2, C to E). We also find significant correlations between background fitness and the variance and skew of the DFE: more-fit segregants have wider DFEs that are skewed toward more-deleterious mutations (Fig. 2, C, F, and G, and fig. S6). These results imply that mutational robustness is negatively correlated with background fitness. This suggests that second-order selection for robustness could be constrained by conflict with first-order selection for fitness.

The DFE is composed of individual mutations. To understand why the shape of the DFE varies between segregants, we examined how the effects of these individual mutations vary. We observe a variety of types of epistasis (Fig. 3 and fig. S7), including nearly constant effects across backgrounds (e.g., PAH1, MME1) and diminishing returns (e.g., SIR3). However, the most frequent pattern (48 cases) is "increasing-cost" epistasis such that the mutation is more deleterious in more-fit segregants (Figs. 3 and 4A). This is the deleterious-mutation analog of diminishing-returns epistasis. Notably, this effect can cross zero: Some mutations are beneficial in the least-fit segregants, neutral in intermediate-fitness segregants, and deleterious in higher-fitness segregants (e.g., RPL16A). This negative correlation is not universal; six mutations exhibit the opposite pattern (e.g., KRI1).

In addition to background fitness, specific genetic loci can influence the fitness effects of individual mutations (e.g., NOP16) (Fig. 3). We used the same procedure as in (*9*, *22*) to identify such quantitative trait loci (QTLs) for each mutation (fig. S8). To quantify how these QTLs and background fitness explain the variation in the fitness effect of each mutation, we fit three linear models to our data (*21*). The "fitness" model includes segregant fitness as the only predictor. The "QTL" model includes only

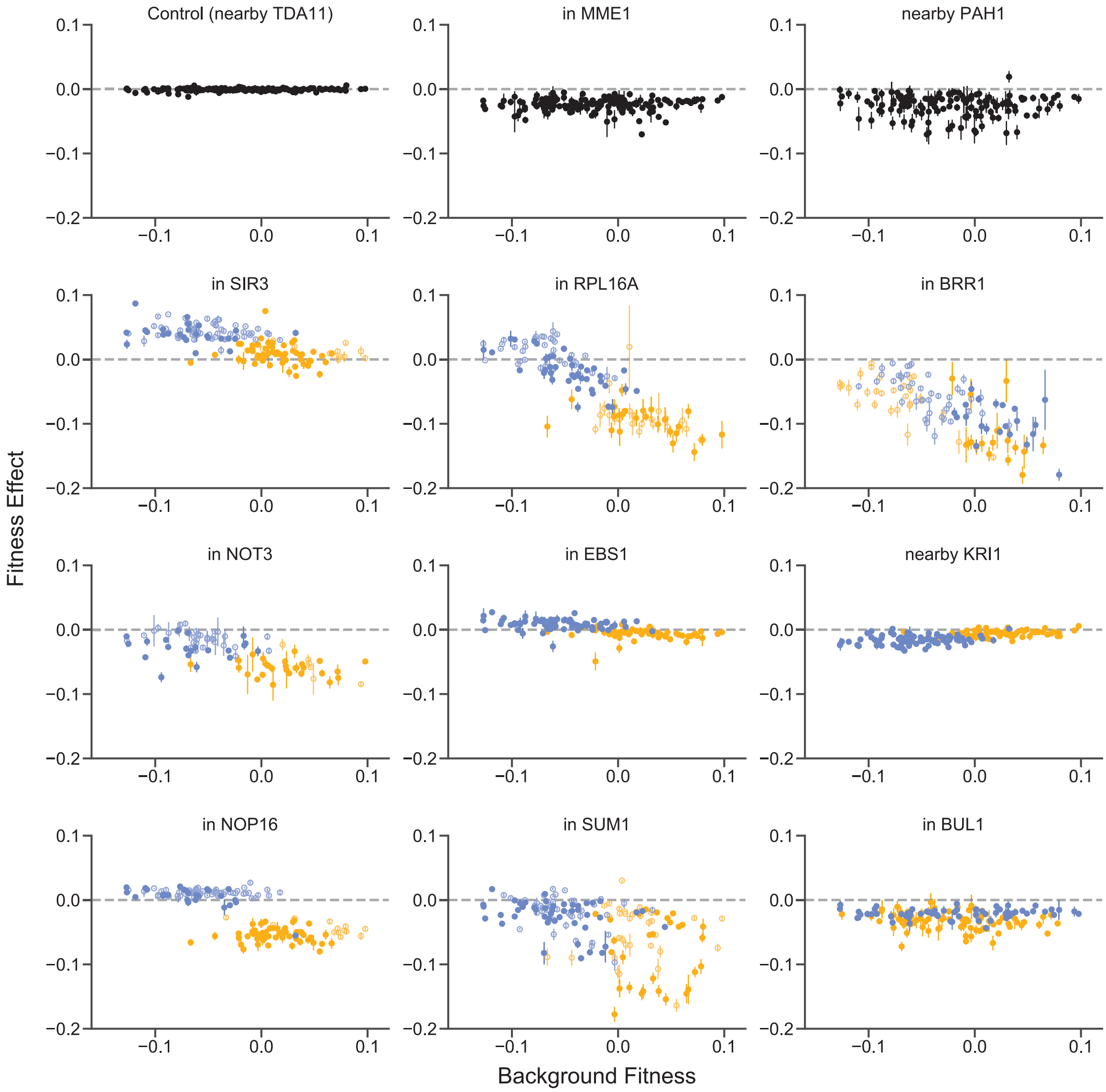


Fig. 3. Patterns of epistasis for individual mutations. Fitness effects of 12 representative insertion mutations are plotted against segregant background fitness. The top-left mutation is one of five putatively neutral insertions used as controls. Allelic state at the largest-effect QTL for the fitness effect of each mutation is shown by yellow (BY) or blue (RM) color; allelic state at the second-largest–effect QTL is shown by closed (BY) or open (RM) symbol. If no significant QTLs were detected, all data points are black. Analogous plots for all insertion mutations are shown in fig. S7. Error bars represent SEs (*21*).

segregant genotype at a small number of QTLs. The "full model" includes both segregant fitness and QTL effects.

We find that at least one of these three models is significantly better than the null model without epistasis for 71 of the 80 mutations for which we have measurements in at least 50 segregants (Fig. 4B). For mutations in which the fitness model has explanatory power (64 cases, 80.0%), most are more deleterious in more-fit backgrounds (58 cases), and most of these are deleterious on average (i.e., exhibit increasing-cost epistasis, 48 cases). QTLs have explanatory power in 60 cases (75.0%), but because many QTLs also affect segregant fitness, QTL and fitness contributions are confounded. To understand which of our models best explains the fitness effects of each mutation while involving as few parameters as possible, we compare the Akaike information criterion (AIC) of each model. The AIC is lowest for the background fitness model for 10 mutations (12.5%), the QTL model for 16 mutations (20%), and the full model for 45 mutations (56.25%). However, even when the background fitness model does not have the lowest AIC, in many cases it explains close to as much variation as the QTL or full model.

One potential explanation for increasing-cost epistasis is that mutations have effects during the saturation phase of our batch culture propagation, and more-fit strains are less robust because they spend longer in saturation. We measured fitness within a single growth cycle and found that this is not the case. Instead, differences in fitness arise almost exclusively during exponential growth (*21*) (figs. S9 to S11).

To understand why faster-growing strains are less robust, we looked for functional similarities among genes disrupted by mutations with similar epistatic patterns [by mutations whose fitness effects were modified by the same QTL (fig. S12) or by mutations with strong fitness-mediated epistasis (*21*)]. We found enrichment in several Gene Ontology (GO) terms among such genes at $P < 0.05$ (though none remain significant after multiple-hypothesis correction) (table S1) (*23*). Most notably, the fitness-mediated epistasis set and the set associated with the most commonly observed QTL are enriched for ribosome- and translation-related functions. This QTL is also the strongest background fitness QTL and includes variants in *KRE33*, a gene involved in small ribosomal subunit assembly (*9*). Metabolic control theory provides a possible link between this functional information and increasing-cost epistasis. Specifically, it predicts that a deleterious mutation in one enzyme will have a weaker effect if other enzymes in any sequential biological pathway are already defective, given that fitness is correlated with metabolic flux through that pathway (*18*, *24*).

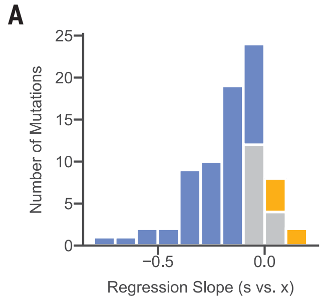
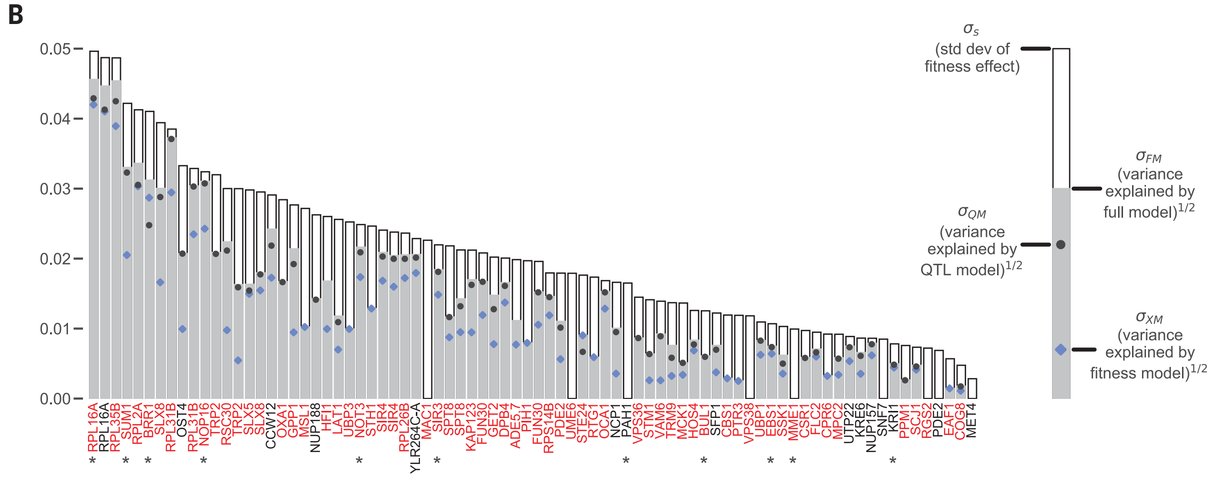


Fig. 4. Genetic determinants of fitness effects. (**A**) Histogram of regression slopes between fitness effect and background fitness for each mutation. Significant negative and positive correlations are shown in blue and yellow, respectively. (**B**) For each mutation, the standard deviation (std dev) of fitness effect across segregants and the square root of the variance explained by each of the three models. For each mutation, the variance explained by models that were not significantly better than a no-epistasis model are not plotted. Mutations shown in red or black are insertions in or near the corresponding gene, respectively; stars indicate the mutations shown in Fig. 3. Only mutations with fitness-effect measurements in at least 50 segregants are shown.



This also applies to sequential pathways in transcription and translation (*25*), so increasing costs epistasis could arise because more-fit segregants have a better-optimized ribosome synthesis or protein synthesis pathway and therefore experience greater costs when deleterious mutations affect these pathways. Although these results are not definitive, they suggest that further generalizations could be drawn from a deeper understanding of the cell-physiological basis of fitness-mediated epistasis.

Our results are limited to the analysis of individual gene disruption mutations in a specific set of yeast strains, and it is possible that other types of mutations (e.g., regulatory changes or second-step mutations) exhibit different patterns. However, to the extent that the patterns of fitness-dependent epistasis observed in this research and in previous work on beneficial mutations (*9*–*15*) hold more broadly and over multiple mutational steps, their net effect is a predictable change in the local properties of the fitness landscape as populations adapt. This local mutational neighborhood becomes less favorable in more-fit genotypes: Uphill steps become flatter or even change to downhill, and many downhill paths become steeper. On such landscapes, first-order selection for high-fitness genotypes conflicts with second-order selection, and populations evolve toward more-fit but less robust and less adaptable genotypes. Thus, even if true fitness peaks exist, populations may never attain them, instead reaching a dynamic balance between beneficial and deleterious mutations (*26*).

REFERENCES AND NOTES

1. S. Wright, *Proc. Sixth Int. Cong. of Genet.* **1**, 356–366 (1932).
2. J. Masel, M. V. Trotter, *Trends Genet.* **26**, 406–414 (2010).
3. R. E. Lenski, J. E. Barrick, C. Ofria, *PLOS Biol.* **4**, e428 (2006).
4. J. L. Payne, A. Wagner, *Nat. Rev. Genet.* **20**, 24–38 (2019).
5. C. O. Wilke, J. L. Wang, C. Ofria, R. E. Lenski, C. Adami, *Nature* **412**, 331–333 (2001).
6. R. J. Woods *et al.*, *Science* **331**, 1433–1436 (2011).
7. I. S. Novella, J. B. Presloid, C. Beech, C. O. Wilke, *J. Virol.* **87**, 4923–4928 (2013).
8. A. S. Lauring, R. Andino, *PLOS Pathog.* **6**, e1001005 (2010).
9. E. R. Jerison *et al.*, *eLife* **6**, e27167 (2017).
10. S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, *Science* **344**, 1519–1522 (2014).
11. A. Wünsche *et al.*, *Nat. Ecol.Evol.* **1**, 61 (2017).
12. M. J. Wiser, N. Ribeck, R. E. Lenski, *Science* **342**, 1364–1367 (2013).
13. A. Couce, O. A. Tenaillon, *Front. Genet.* **6**, 99 (2015).
14. H.-H. Chou, H.-C. Chiu, N. F. Delaney, D. Segrè, C. J. Marx, *Science* **332**, 1190–1192 (2011).
15. A. I. Khan, D. M. Dinh, D. Schneider, R. E. Lenski, T. F. Cooper, *Science* **332**, 1193–1196 (2011).
16. S. F. Elena, R. E. Lenski, *Nature* **390**, 395–398 (1997).
17. R. Sanjuán, A. Moya, S. F. Elena, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15376–15379 (2004).
18. L. Jasnos, R. Korona, *Nat. Genet.* **39**, 550–554 (2007).
19. M. Costanzo *et al.*, *Science* **327**, 425–431 (2010).
20. A. Kumar *et al.*, *Genome Res.* **14** (10a), 1975–1986 (2004).
21. Materials and methods and supplementary text are available as supplementary materials.
22. J. S. Bloom, I. M. Ehrenreich, W. T. Loo, T.-L. V. Lite, L. Kruglyak, *Nature* **494**, 234–237 (2013).
23. D. V. Klopfenstein *et al.*, *Sci. Rep.* **8**, 10872 (2018).
24. E. Szathmáry, *Genetics* **133**, 127–132 (1993).
25. R. C. MacLean, *J. Evol. Biol.* **23**, 488–493 (2010).
26. D. P. Rice, B. H. Good, M. M. Desai, *Genetics* **200**, 321–329 (2015).
27. M. S. Johnson, TnSeq Analysis Pipeline Publication Archive, Zenodo (2019); 10.5281/zenodo.3402230.

ACKNOWLEDGMENTS

We thank E. Jerison, L. Rast, A. Nguyen-Ba, J. Yodh, G. Wildenberg, and members of the Desai lab for experimental assistance and/or comments on the manuscript. **Funding:** This work was supported by an NSF Graduate Research Fellowship (to M.S.J.), a BWF Career Award at the Scientific Interface (grant 1010719.01), the Alfred P. Sloan Foundation (grant FG-2017-9227), the Hellman Foundation, the Simons Foundation (grant 376196), the NSF (DEB-1655960), and the NIH (GM104239). Computational work was performed on the Odyssey cluster supported by the Research Computing Group at Harvard University. **Author contributions:** Conceptualization: S.K. and M.M.D.; experimental design: M.S.J., A.M., S.K., and M.M.D; methods development: M.S.J. and A.M.; experiments: M.S.J.; analysis: M.S.J. and S.K.; writing: M.S.J., S.K., and M.M.D. **Competing interests:** None declared. **Data and materials availability:** Data described in the paper are presented in the supplementary materials. Raw sequencing data are publicly available at the NCBI Sequence Read Archive (accession no. SRP216610), and all analysis code is available from Zenodo (*27*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/490/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S6
References (*28*–*40*)
Data File Descriptions
Data Files S1 to S3 (Excel format)

18 June 2019; accepted 11 September 2019
10.1126/science.aay4199

IMMUNOLOGY

Microbial metabolites control the thymic development of mucosal-associated invariant T cells

François Legoux[1]*, Déborah Bellet[2], Celine Daviaud[1,3], Yara El Morr[1], Aurelie Darbois[1], Kristina Niort[4], Emanuele Procopio[1], Marion Salou[1], Jules Gilet[1], Bernhard Ryffel[5], Aurélie Balvay[2], Anne Foussier[2], Manal Sarkis[1,6], Ahmed El Marjou[4], Frederic Schmidt[6], Sylvie Rabot[2], Olivier Lantz[1,7,8]*

How the microbiota modulate immune functions remains poorly understood. Mucosal-associated invariant T (MAIT) cells are implicated in mucosal homeostasis and absent in germ-free mice. Here, we show that commensal bacteria govern murine MAIT intrathymic development, as MAIT cells did not recirculate to the thymus. MAIT development required *RibD* expression in bacteria, indicating that production of the MAIT antigen 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU) was necessary. 5-OP-RU rapidly traveled from mucosal surfaces to the thymus, where it was captured by the major histocompatibility complex class Ib molecule MR1. This led to increased numbers of the earliest MAIT precursors and the expansion of more mature receptor-related, orphan receptor γt–positive MAIT cells. Thus, a microbiota-derived metabolite controls the development of mucosally targeted T cells in a process blurring the distinction between exogenous antigens and self-antigens.

Mucosal-associated invariant T (MAIT) cells are evolutionarily conserved T cells that recognize vitamin B2 precursor derivatives [5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU)] presented by the major histocompatibility complex class Ib molecule MR1 (*1*). These derivatives are produced by most bacteria and yeasts but not by animal cells (*2*). MAIT cells have been implicated in human pathologies associated with microbiotal dysbiosis (*3*–*6*). In mice, mature, receptor-related orphan receptor γt–positive (RORγt$^+$) MAIT cells play beneficial roles in maintaining intestinal homeostasis (*5*, *7*). Germ-free (GF) mice lack MAIT cells (*8*, *9*), but the underlying mechanisms are unknown. Here, we report that vitamin B2 metabolites are directly transferred into the thymus and presented to thymocytes, driving the intrathymic expansion of RORγt$^+$ MAIT cells. Thus, commensal bacteria influence barrier homeostasis by controlling the thymic production of T cells targeted to mucosae.

To investigate the role of the microbiota in MAIT development, we compared thymic MAIT frequencies of mice under specific pathogen–free (SPF) and GF conditions (*10*). As recently described (*8*), the number of MR1:5-OP-RU tetramer$^+$ (MAIT) cells was reduced in the thymus (Fig. 1A), as well as in the spleen (Fig. 1B), lungs (Fig. 1C), and colon (fig. S1A), of GF mice. By contrast, invariant natural killer T (iNKT) cells, a T cell subset with different specificity but parallel development (*11*, *12*), were not reduced in GF mice (fig. S1B) (*13*). We then characterized the developmental defect of MAIT cells as described previously (*8*). HSA$^+$CD44$^-$ (Heat-stable antigen–positive, CD44-negative) immature MAIT cells (*8*) were reduced in GF mice (Fig. 1, D and E). A few cells reached the HSA$^-$CD44$^+$ most mature stage (Fig. 1, D and E) and expressed the lineage-defining transcription factor PLZF (promyelocytic leukemia zinc finger) at normal levels (fig. S1C). These MAIT cells may be selected by empty MR1 or by MR1 loaded with unknown self-ligands. We compared the sublineage choice of mature MAIT cells in SPF and GF mice using CD122 and CD138 to identify the T-bet$^+$ MAIT1 and RORγt$^+$ MAIT17 subsets, respectively (fig. S1D). MAIT17 cells were decreased in GF mice in both the thymus (Fig. 1, F and G, and fig. S1F) and spleen (fig. S1E). Accordingly, MAIT cells in GF mice were skewed toward interferon (IFN)-γ production with reduced interleukin (IL)-17 production (Fig. 1H).

To directly assess the role of microbiota on MAIT development, adult GF mice were cohoused with SPF mice. Upon microbial colonization, MAIT frequencies increased first in the thymus, then in the spleen and lungs (Fig. 1, A to C), without reaching in the periphery the levels of mice colonized at birth. In the thymus, numbers of both immature HSA$^+$ and mature CD44$^+$ MAIT cells increased and reached SPF levels 2 weeks after microbial colonization (Fig. 1D). The increase in mature thymic MAIT numbers was accounted for by the expansion of MAIT17 cells, whereas MAIT1 numbers remained unchanged (Fig. 1E). Thus, commensal microbes are necessary for the complete thymic maturation of MAIT17 cells.

MAIT cells may expand in the periphery and circulate back to the thymus, thereby explaining the reduced frequency of mature MAIT cells in GF thymi. We used parabiosis of B6-MAIT^{Cast} mice, which exhibit increased frequency of MAIT cells compared with C57BL/6 mice (fig. S2A) (*14*), to address this possibility (fig. S2B). Only 2.5 ± 1% (SEM) of mature thymic MAIT cells originated from the other parabiont, indicating minimal migration from the periphery (Fig. 2A). Few MAIT cells also migrated into $Mr1^{-/-}$ parabionts (fig. S2C). Accordingly, MAIT17 cells were highly proliferative (Ki67$^+$) within the thymus but not in the spleen (Fig. 2B). Finally, MAIT frequency was unaffected in the thymus, spleen, and lungs of B6-MAIT^{Cast} $Jh^{-/-}$ mice (fig. S2, D and E), which lack B cells. Thus, B cells are not required for MAIT development.

Given the similarities in the thymic differentiation process of iNKT and MAIT cells (*12*, *15*, *16*), their discordant dependency on microbiota suggested a role for T cell receptor (TCR) specificity and antigen (Ag) availability in the thymus. Therefore, we measured expression levels of Nur77, a transcription factor correlated with TCR signaling strength (*17*). Nur77 was expressed in MAIT17 cells in the thymus but not in the spleen or in the MAIT1 cells of either organ (Fig. 2C). Thus, MAIT17 cells receive strong TCR signals in the thymus. We then investigated whether the phosphatidylinositol 3-kinase (PI3K) pathway, which controls proliferation in lymphocytes (*18*), was involved in MAIT development. Deletion of phosphatase and tensin homolog (PTEN), a negative regulator of PI3K activity, had no effect on the number of MAIT1 cells, yet caused a 53-fold increase in MAIT17 numbers in the thymus but not the spleen (Fig. 2D). This suggested that the PI3K pathway was active in thymic MAIT17 cells. The discrepancy between MAIT1 and MAIT17 behaviors may result from the expression of several inhibitory receptors, such as the TCR inhibitor *Ptpn22*, by MAIT1 cells (fig. S2F) (*12*). The ability of both MAIT subsets to signal through the TCR was evaluated in vitro. Although the majority of MAIT17 cells expressed Nur77 in response to 5-OP-RU, only a fraction of MAIT1 cells did (fig. S2G), indicating that TCR signaling was inhibited in this subset. Thus, only MAIT17 cells receive strong TCR signals and proliferate in a PI3K-dependent manner in the thymus.

In GF mice, MAIT17 cells expressed lower levels of Ki67 (Fig. 2E), as well as lower levels of Nur77 (Fig. 2F), indicative of reduced TCR signaling. Nur77 expression was lower in immature HSA$^+$ cells from GF mice, which indicates that the microbiota contributed to TCR stimulation in these cells as well. We recently reported increasing TCR Vβ8 expression in

[1]INSERM U932, PSL University, Institut Curie, Paris 75005, France. [2]Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas 78350, France. [3]Animal Facility, Institut Curie, Paris 75005, France. [4]Recombinant Protein Facility, Institut Curie, Paris 75005, France. [5]CNRS UMR7355, Université d'Orléans, Orléans 45067, France. [6]CNRS UMR3666, INSERM U1143, PSL University, Institut Curie, Paris 75005, France. [7]Laboratoire d'Immunologie Clinique, Institut Curie, Paris 75005, France. [8]Centre d'Investigation Clinique en Biothérapie Gustave-Roussy Institut Curie (CIC-BT1428), Institut Curie, Paris 75005, France.
***Corresponding author. Email: olivier.lantz@curie.fr (O.L.); francois.legoux@curie.fr (F.L.)**

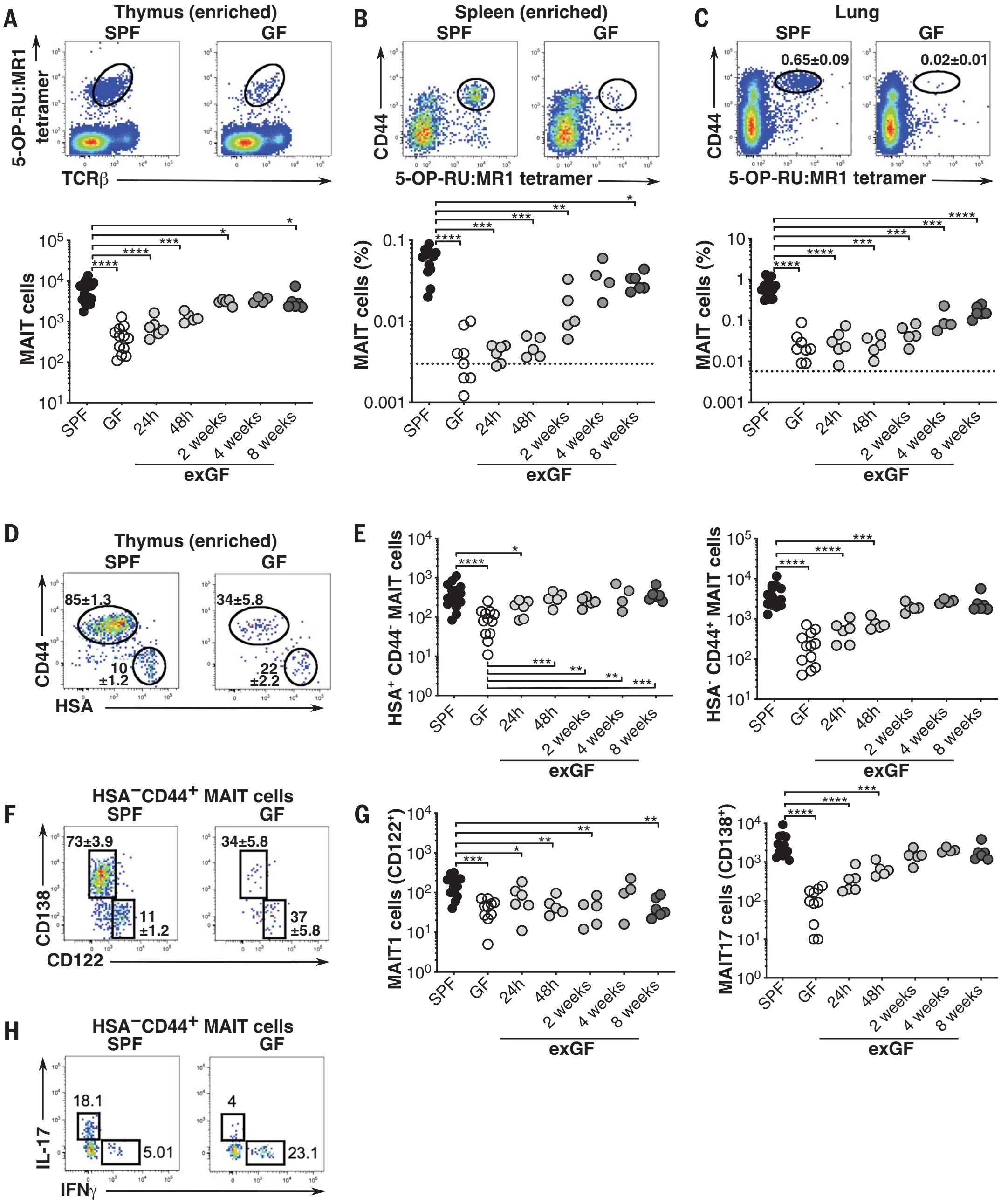


Fig. 1. Rapid development of MAIT17 cells upon colonization with commensal microbes. (**A** to **C**) Top: Representative flow cytometry of thymocytes (A), splenocytes (B), and lung cells (C) from SPF and GF B6 mice. Data are shown as mean ± SEM; $N = 3$ independent experiments; $n \geq 1$ biological replicates per experiment. exGF, GF mice co-housed with SPF mice. Bottom: MAIT cell numbers in thymus (A) or MAIT cell frequencies in spleen (B) and lungs (C) of the indicated mice. Dotted lines represent the limit of detection as defined by the mean frequency of 6-FP:MR1 tetramer$^+$ events. (**D**) Representative HSA and CD44 expression in MAIT thymocytes. Data are shown as mean ± SEM; $N = 4$ experiments, $n \geq 1$ replicates. (**E**) Numbers of HSA$^+$CD44$^-$ and HSA$^-$CD44$^+$ MAIT thymocytes in the indicated mice. (**F**) Representative CD122 and CD138 expression in HSA$^-$CD44$^+$ MAIT thymocytes. Data are shown as mean ± SEM; $N = 3$ experiments, $n \geq 1$ replicates. (**G**) Numbers of MAIT1 and MAIT17 thymocytes in the indicated mice. (**H**) IL-17 and IFN-γ expression in HSA$^-$CD44$^+$ MAIT thymocytes after phorbol 12-myristate 13-acetate and ionomycin stimulation. A total of eight to nine thymi were pooled together; $N = 2$ experiments, $n = 1$ replicate. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ by Mann–Whitney *U* tests.

MAIT cells during thymic maturation (*19*). In GF mice, Vβ8 expression did not increase with MAIT maturation (Fig. 2G), consistent with a role for microbiota-derived Ags in shaping the thymic development of MAIT cells.

The ability of a bacterial strain to trigger MAIT TCRs strictly correlates with the ability to produce an intermediary compound of the vitamin B2 biosynthesis pathway, 5-amino-ribityl uracil (5-A-RU), which reacts with methylglyoxal to generate 5-OP-RU (*20*–*22*). Because microbial metabolites may circulate in the body (*23*, *24*), and colonization with commensal microbes results in the expansion of Nur77^{hi} MAIT cells in the thymus, we investigated whether 5-OP-RU could be captured and presented by thymic cells. In vitro incubation with synthetic 5-OP-RU, but not 5-A-RU, resulted in the up-regulation of MR1 at the cell surface of thymocytes, indicating efficient capture by intracellular MR1 (fig. S3A). Using an in vitro assay, we compared the ability of cells from different tissues to capture and present synthetic 5-OP-RU. Thymic cells were 50 times more efficient than splenocytes at activating 5-OP-RU:MR1–specific TCR-Tg reporter cells in vitro (Fig. 3A and fig. S3B). Thymic cells from GF mice also captured and presented synthetic 5-OP-RU to reporter cells in vitro (fig. S3C). By contrast, thymocytes were no better than splenocytes at capturing and presenting the iNKT ligand α-galactosylceramide (αGalCer) to αGalCer:CD1d–specific TCR Tg cells (fig. S3D).

To study the in vivo dynamics of 5-OP-RU, we injected decreasing doses of 5-OP-RU (or αGalCer control) intraperitoneally into wild-type or *Mr1*$^{-/-}$ mice. In mice administered with αGalCer, cells from the spleen and liver, but not the thymus, activated αGalCer:CD1d–specific reporter cells in vitro (fig. S3D). Thus, thymocytes did not capture circulating αGalCer, confirming the current view of the thymus being hermetic to circulatory Ags. By contrast, in mice receiving 5-OP-RU, thymic cells were more efficient at activating 5-OP-RU:MR1–specific reporter cells in vitro than were cells isolated from the periphery (Fig. 3B). Thus,

systemic 5-OP-RU was efficiently captured and presented by thymic cells. The presentation of 5-OP-RU was MR1 dependent as thymic cells from $Mr1^{-/-}$ mice did not activate reporter cells. Neither 5-A-RU nor the nonstimulatory MR1 ligand acetyl-6-formylpterin (Ac6FP) activated reporter cells after injection (fig. S3E). Accordingly, injection of 5-OP-RU, but not 5-A-RU, triggered Ki67 expression in endogenous MAIT17 cells (Fig. 3C). Single-positive thymocytes, double-positive (DP) thymocytes, dendritic cells, and medullary and cortical thymic epithelial cells isolated as outlined in fig. S3F were all able to activate reporter cells upon in vitro incubation with 5-OP-RU. However,

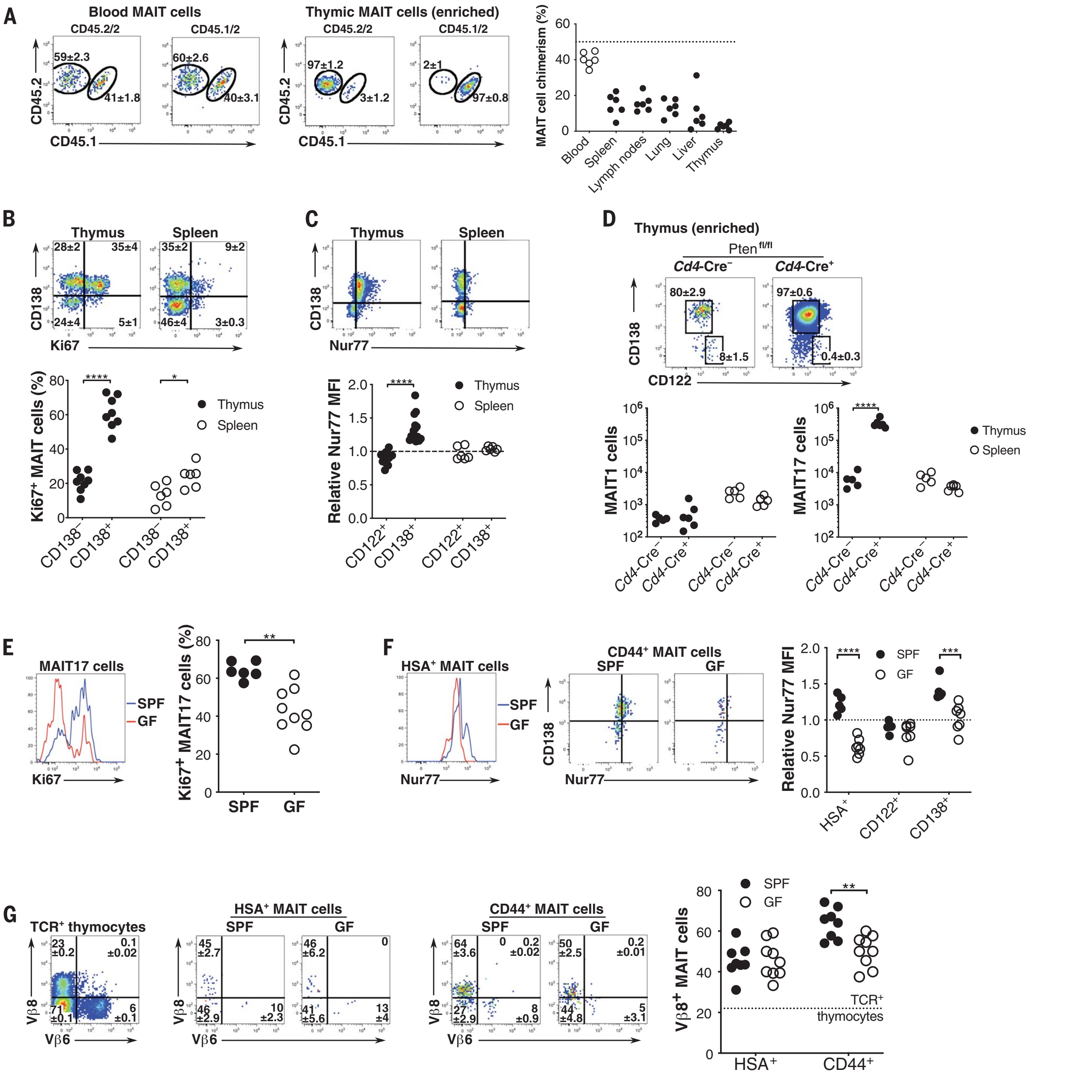


Fig. 2. MAIT17 cells receive TCR signals and proliferate in the thymus. (**A**) Representative flow cytometry of MAIT cells in the blood (left) and thymus (center) of CD45.1/2 and CD45.2/2 congenic B6-MAIT^{Cast} parabiotic mice. Data are shown as mean ± SEM; N = 3 independent experiments, n = 2 biological replicates per experiment. Right: Chimerism of MAIT cells in the indicated tissues. (**B**) Top: Representative flow cytometry of HSA^{-}CD44^{+} MAIT cells from thymus and spleen. Data are shown as mean ± SEM; N = 4 experiments, $n \geq 1$ replicates. Bottom: Proportion of Ki67^{+} cells in the indicated MAIT subset. (**C**) Top: Representative flow cytometry of HSA^{-}CD44^{+} MAIT cells from thymus and spleen; N = 3 experiments, $n \geq 1$ replicates. Bottom: Mean fluorescence intensity (MFI) of Nur77 staining in the indicated MAIT subset relative to CD4^{-}CD8^{+} cells from the same sample. (**D**) Top: Representative flow cytometry of HSA^{-}CD44^{+} MAIT cells from thymus of the indicated mice. Data are shown as mean ± SEM; N = 3 experiments, $n \geq 1$ replicates. Bottom: Number of MAIT1 and MAIT17 cells in the thymus or spleen of the indicated mice. (**E** to **G**) Left: Representative flow cytometry of the indicated MAIT subset from the thymus of SPF or GF mice; $N \geq 2$ experiments, $n \geq 2$ replicates. Right: Proportion of Ki67^{+} MAIT17 cells (E), MFI of Nur77 staining relative to CD4^{-}CD8^{+} cells (F), and proportion of TCR Vβ8^{+} MAIT cells (G) in the thymus of SPF or GF mice. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ by Student's t tests.

DP thymocytes were the most potent antigen-presenting cells (APCs) (Fig. 3D). After 5-OP-RU injection, DP thymocytes, dendritic cells, and epithelial cells all captured and presented 5-OP-RU in an MR1-dependent fashion (Fig. 3E).

We then investigated whether 5-A-RU or 5-OP-RU could reach the thymus from outside of the body. When painted onto the intact skin of mice, no αGalCer was detected in any tissue (fig. S3G). By contrast, only thymic cells from mice painted with 5-OP-RU, but not 5-A-RU, activated reporter cells (Fig. 3, F and G). Thus, some 5-OP-RU from the skin was captured in the thymus but not by other sampled tissues. The cutaneous application of 5-OP-RU, but not 5-A-RU, also activated endogenous thymic MAIT cells, as well as splenic MAIT cells to a lesser extent (fig. S3H). 5-OP-RU was detected in the thymus 1 hour after cutaneous application or after oral gavage (Fig. 3H). Thus, exogenous 5-OP-RU rapidly reaches the thymus and is presented to MAIT thymocytes by various APCs.

MAIT cell development may be directly dependent upon the production of 5-OP-RU by commensal bacteria. To control for microbial factors not associated with 5-OP-RU production, we used *Escherichia coli* strains with genetic deletion of vitamin B2 enzymes either upstream (Δ*RibD*) or downstream (Δ*RibE*) of 5-A-RU production (*22*) (Fig. 4A).

GF mice were monocolonized with either Δ*RibD* or Δ*RibE E. coli*. MAIT frequency remained low in thymus and lungs of mice colonized with Δ*RibD* bacteria. By contrast, colonization with Δ*RibE* bacteria induced MAIT development and migration into lungs (Fig. 4B). 5-OP-RU production by bacteria was required for thymic expansion of MAIT17 cells (Fig. 4, C and D). 5-OP-RU production also controlled increased numbers of immature HSA⁺ MAIT cells (Fig. 4C). This may reflect either positive selection of new MAIT cells or the expansion of preselected HSA⁺ MAIT cells. These results were confirmed with the wild-type commensal bacterial species *Enterococcus hirae* (vitamin B2 proficient) and *Enterococcus faecalis* (vitamin B2 deficient) (*25*) (fig. S4, A to E). Thus, 5-OP-RU production by commensal bacteria is required for the complete maturation of thymic MAIT cells.

To test whether 5-A-RU-derived Ags are sufficient to induce MAIT development, SPF and GF mice were injected with various doses

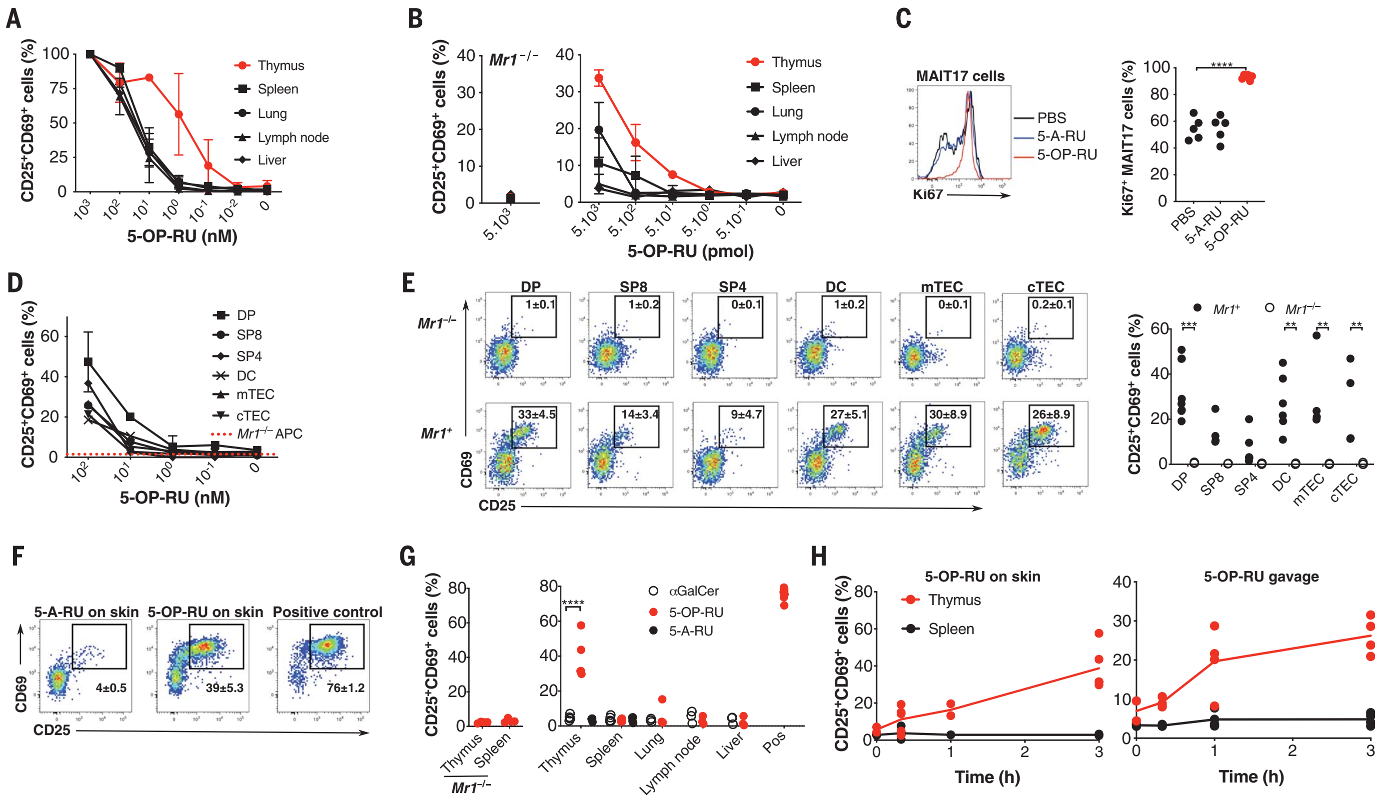


Fig. 3. Vitamin B2 metabolite 5-OP-RU can cross mucosal barriers and reach the thymus for presentation to MAIT cells. (**A**) Relative activation of reporter splenocytes incubated with total cells from the indicated tissues pulsed with 5-OP-RU. Data are shown as mean ± SEM; $N = 3$ independent experiments, $n = 1$ biological replicate per experiment. (**B**) Activation of reporter splenocytes incubated with total cells from the indicated tissues harvested 3 hours after intraperitoneal injection of 5-OP-RU. Data are shown as mean ± SEM; $N = 2$ experiments, $n = 1$ replicate. (**C**) Left: Representative flow cytometry of MAIT17 thymocytes 24 hours after intraperitoneal injection of the indicated molecule (5 nmol); $N = 2$ experiments, $n \geq 2$ replicates. Right: Proportion of Ki67⁺ MAIT17 cells in the indicated condition. PBS, phosphate-buffered saline. (**D**) Activation of reporter splenocytes incubated with thymic APCs isolated by fluorescence-activated cell sorting and pulsed with 5-OP-RU. Data are shown as mean ± SEM; $N = 2$ experiments. DC, dendritic cell; mTEC, medullary thymic epithelial cell; cTEC, cortical thymic epithelial cell. (**E**) Left: Representative flow cytometry of reporter splenocytes incubated with thymic APCs from *Mr1*$^{-/-}$ or *Mr1*$^{+}$ mice previously injected with 5-OP-RU (5 nmol). Data are shown as mean ± SEM; $N = 3$ experiments, $n \geq 1$ replicates. Right: Activation of reporter splenocytes incubated with the indicated APCs. (**F**) Representative flow cytometry of reporter splenocytes incubated with total thymocytes from mice to which αGalCer (5 nmol) or 5-OP-RU (5 nmol) was painted on intact skin 3 hours before tissue harvest. As a positive control, thymocytes were incubated with 5-OP-RU (1 μM). Data are shown as mean ± SEM; $N = 3$ experiments, $n \geq 1$ replicates. (**G**) Summary of the data obtained as in (F). (**H**) Activation of reporter splenocytes incubated with thymocytes or splenocytes from mice having received 5-OP-RU (5 nmol) either on the skin (left panel) or through oral gavage (right panel) for the indicated times before tissue harvest; $N = 2$ experiments, $n \geq 1$ replicates. $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$ by Student's *t* tests.

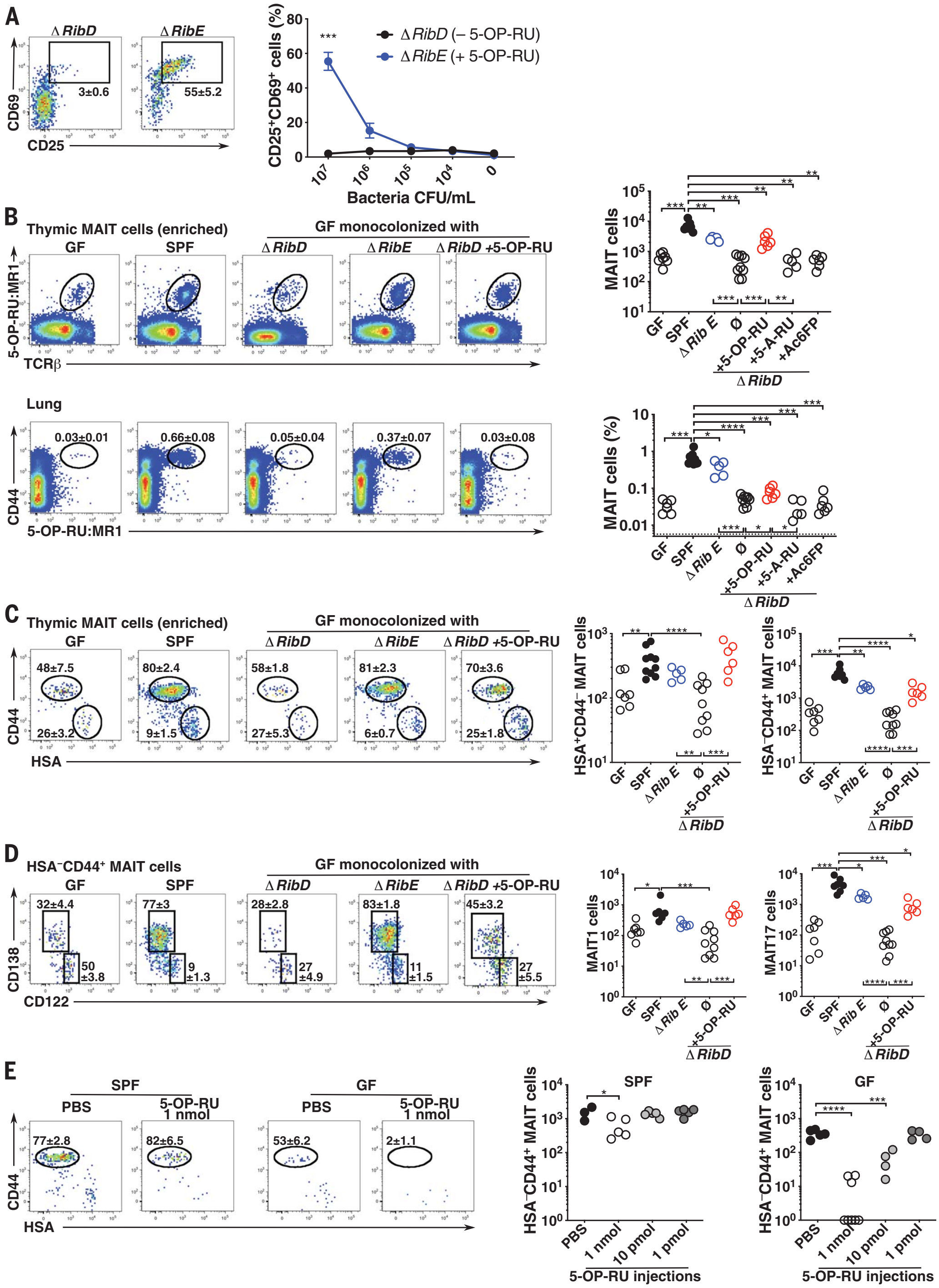


Fig. 4. 5-OP-RU produced by commensal bacteria controls MAIT cell development in the thymus. (**A**) Left: Representative flow cytometry of reporter splenocytes incubated with thymocytes previously cocultured with 10^7 colony-forming units/ml of the indicated strain of *E. coli*. Data are shown as mean ± SEM; *N* = 2 independent experiments, *n* = 2 biological replicates per experiment. Right: Summary of the data. (**B**) Left: Representative flow cytometry of cells from the thymus (top) and lungs (bottom) of the indicated mice; *N* ≥ 2 experiments, *n* ≥ 2 replicates. Right: MAIT cell numbers in the thymus (top) and MAIT frequencies in the lungs (bottom) of the indicated mice. (**C**) Left: Representative HSA and CD44 expression in MAIT thymocytes. Data are shown as mean ± SEM; *N* ≥ 2 experiments, *n* ≥ 2 replicates. Right: Numbers of HSA^+CD44^- and HSA^-CD44^+ MAIT thymocytes in the indicated mice. (**D**) Left: Representative CD122 and CD138 expression in HSA^-CD44^+ MAIT thymocytes in the indicated mice. Data are shown as mean ± SEM; *N* ≥ 2 experiments, *n* ≥ 2 replicates. Right: Numbers of MAIT1 and MAIT17 cells in the indicated mice. (**E**) Left: Representative flow cytometry of MAIT thymocytes in mice injected with PBS or with 1 nmol of 5-OP-RU for 2 weeks. Data are shown as mean ± SEM; *N* = 2 experiments, *n* ≥ 1 replicates. Right: Numbers of HSA^-CD44^+ MAIT thymocytes in the indicated conditions. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ by Student's *t* tests (A) or by Mann–Whitney *U* tests (B to E).

of 5-OP-RU. These injections induced a dose-dependent depletion of thymic MAIT cells in GF mice (Fig. 4E), likely a result of negative selection. MAIT cells were also depleted from the thymus of SPF mice but required higher doses of 5-OP-RU, suggesting that additional factors induced by the microbiota promoted MAIT survival. MAIT development was not impaired in $Myd88^{-/-}$ or $Tlr3^{-/-}$ mice (fig. S4F), ruling out a role for Toll-like receptor–dependent microbial sensing. Because myeloid differentiation factor 88 is necessary for signaling through IL-1 receptor family members, IL-1, IL-18, and IL-33 are therefore dispensable for MAIT development. Aryl hydrocarbon receptor, IFN-γ, and IFN-γ receptor were also dispensable for MAIT development (fig. S4F), suggesting the involvement of other microbiota-induced mediators in MAIT expansion and maintenance.

To test whether 5-OP-RU, together with an unknown bacterially induced factor(s), controlled MAIT maturation, low doses of 5-OP-RU (10 pmol) were injected into GF mice (Fig. 4E) and GF mice monocolonized with Δ*RibD* bacteria (Fig. 4, B to D). Injections of 5-OP-RU resulted in increased MAIT numbers in the thymus (Fig. 4B), in association with increased HSA^+ and $CD44^+$ MAIT numbers (Fig. 4C). The injection of Ac6FP or 5-A-RU had no effect on MAIT cells (Fig. 4B), indicating that methylglyoxal from the mouse cannot react in vivo with 5-A-RU to form agonist ligands. Thus,

bacterial metabolites such as 5-OP-RU drive the thymic expansion of MAIT17 cells.

The results of this study help to explain how the microbiota can influence the host at distant sites through the production of metabolites.

REFERENCES AND NOTES

1. K. Franciszkiewicz *et al*., *Immunol. Rev.* **272**, 120–138 (2016).
2. S. Mondot, P. Boudinot, O. Lantz, *Immunogenetics* **68**, 537–548 (2016).
3. I. Magalhaes *et al*., *J. Clin. Invest.* **125**, 1752–1762 (2015).
4. Y. Miyazaki, S. Miyake, A. Chiba, O. Lantz, T. Yamamura, *Int. Immunol.* **23**, 529–535 (2011).
5. O. Rouxel *et al*., *Nat. Immunol.* **18**, 1321–1331 (2017).
6. N. E. Serriari *et al*., *Clin. Exp. Immunol.* **176**, 266–274 (2014).
7. A. Varelias *et al*., *J. Clin. Invest.* **128**, 1919–1936 (2018).
8. H. F. Koay *et al*., *Nat. Immunol.* **17**, 1300–1311 (2016).
9. E. Treiner *et al*., *Nature* **422**, 164–169 (2003).
10. Materials and methods are available as supplementary materials.
11. F. Legoux, M. Salou, O. Lantz, *Annu. Rev. Cell Dev. Biol.* **33**, 511–535 (2017).
12. M. Salou *et al*., *J. Exp. Med.* **216**, 133–151 (2019).
13. S. H. Park, K. Benlagha, D. Lee, E. Balish, A. Bendelac, *Eur. J. Immunol.* **30**, 620–625 (2000).
14. Y. Cui *et al*., *J. Clin. Invest.* **125**, 4171–4185 (2015).
15. A. Bendelac, *J. Exp. Med.* **182**, 2091–2096 (1995).
16. N. Seach *et al*., *J. Immunol.* **191**, 6002–6009 (2013).
17. A. E. Moran *et al*., *J. Exp. Med.* **208**, 1279–1289 (2011).
18. K. Okkenhaug, *Annu. Rev. Immunol.* **31**, 675–704 (2013).
19. F. Legoux *et al*., *Nat. Immunol.* **20**, 1244–1255 (2019).
20. A. J. Corbett *et al*., *Nature* **509**, 361–365 (2014).
21. L. Le Bourhis *et al*., *Nat. Immunol.* **11**, 701–708 (2010).
22. C. Soudais *et al*., *J. Immunol.* **194**, 4641–4649 (2015).
23. M. Gomez de Agüero *et al*., *Science* **351**, 1296–1302 (2016).
24. Y. Uchimura *et al*., *Immunity* **49**, 545–559.e5 (2018).
25. M. Kanehisa, S. Goto, *Nucleic Acids Res.* **28**, 27–30 (2000).

ACKNOWLEDGMENTS

We thank V. Dangles-Marie, M. Garcia, I. Grandjean, the mouse facility technicians, and the flow cytometry core at Institut Curie. We also acknowledge G. Eberl, S. Latour, and A. Lehuen for mice; P. Serror for bacteria; and N. Manel and S. Amigorena for discussions and for reviewing the manuscript. We thank the NIH tetramer core facility (Emory University) for providing CD1d and MR1 tetramers. The MR1:5-OP-RU tetramer technology was developed jointly by J. McCluskey, J. Rossjohn, and D. Fairlie, and the material was produced by the NIH Tetramer Core Facility as permitted to be distributed by the University of Melbourne. **Funding:** F.L. was supported by a Marie-Skłodowska Curie individual fellowship (706353) from the European Commission (H2020). This work was supported by the Institut National de la Santé et de la Recherche Médicale, Institut Curie, and Agence Nationale de la Recherche (ANR) [Blanc (neoMAIT, MAIT, and diabMAIT) and Labex DCBIOL]. O.L.'s group is supported by the Equipe Labellisée de la Ligue Contre le Cancer. **Author contributions:** F.L. designed the project and performed experiments, analyzed data, and wrote the paper. C.D. designed and performed experiments. Y.E.M. and A.D. performed experiments and analyzed data. D.B., E.P., A.F., and A.B. performed experiments. M.Sal., J.G., and S.R. provided technical support and advice. A.E.M., K.N., M.Sar., and F.S. developed and provided reagents. B.R. provided mice. O.L. supervised the project, designed experiments, and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All experimental data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/494/suppl/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References (*26*–*32*)

View/request a protocol for this paper from *Bio-protocol*.

4 December 2018; resubmitted 15 May 2019
Accepted 14 August 2019
Published online 29 August 2019
10.1126/science.aaw2719

INFLUENZA

Broadly protective human antibodies that target the active site of influenza virus neuraminidase

Daniel Stadlbauer[1,2]*, Xueyong Zhu[3]*, Meagan McMahon[1], Jackson S. Turner[4], Teddy J. Wohlbold[1,5,6], Aaron J. Schmitz[4], Shirin Strohmeier[1], Wenli Yu[3], Raffael Nachbagauer[1], Philip A. Mudd[7], Ian A. Wilson[3,8]†, Ali H. Ellebedy[4]†, Florian Krammer[1]†

Better vaccines against influenza virus are urgently needed to provide broader protection against diverse strains, subtypes, and types. Such efforts are assisted by the identification of novel broadly neutralizing epitopes targeted by protective antibodies. Influenza vaccine development has largely focused on the hemagglutinin, but the other major surface antigen, the neuraminidase, has reemerged as a potential target for universal vaccines. We describe three human monoclonal antibodies isolated from an H3N2-infected donor that bind with exceptional breadth to multiple different influenza A and B virus neuraminidases. These antibodies neutralize the virus, mediate effector functions, are broadly protective in vivo, and inhibit neuraminidase activity by directly binding to the active site. Structural and functional characterization of these antibodies will inform the development of neuraminidase-based universal vaccines against influenza virus.

Seasonal influenza virus infections cause substantial global morbidity and mortality annually. In addition, pandemics occur at irregular and unpredictable intervals and can claim millions of lives. Current seasonal influenza vaccines induce narrow, strain-specific immune responses, and their effectiveness varies according to how well they match the circulating virus strains (*1*). These vaccines also do not protect against new pandemic or emerging viruses. Therefore, broadly protective or universal influenza virus vaccines are urgently needed (*2*). Current vaccines are designed to induce antibody responses against the major surface glycoprotein of the virus, the hemagglutinin (HA) (*3*). Antibodies to HA, specifically to its immunodominant globular head domain, can block virus binding to its sialic acid receptor, but such activity is often restricted to the vaccine strain. Antibodies to the less dominant HA stalk have been shown to have much greater breadth within and across influenza A groups 1 and 2, and rarely also to influenza B viruses (*4*–*7*).

The second major virus surface glycoprotein is the neuraminidase (NA), which cleaves the terminal sialic acid from N-linked glycans that also act as receptors (*8*). The enzymatic activity of NA is important for releasing incoming viruses trapped by glycans of natural defense proteins on mucosal surfaces and for the release of nascent viruses as they bud from infected cells. Anti-NA antibodies can block interactions between NA and its substrate (*9*, *10*). NA monoclonal antibodies (mAbs) and NA vaccination protect against influenza virus lethal challenge in animal models (*9*–*15*). Furthermore, NA vaccination in guinea pigs can prevent virus transmission (*16*). Most important, anti-NA antibodies independently correlate with protection from infection in field studies, as well as in human challenge studies (*17*–*19*).

The dominant antigenic sites on HA mutate at a high rate because of immune pressure, the error rate of the polymerase, and the high plasticity of its globular head domain (*20*). NA, on the other hand, exhibits a slower drift that is generally discordant with that of HA (*21*, *22*). Therefore, antibody responses against NA typically show broader cross-reactivity; however, this breadth has been assumed to be limited to the respective subtype (N1 to N9 for influenza A viruses) (*8*). Here, we report three clonally related mAbs derived from plasmablasts isolated from an H3N2-infected individual that show broad heterosubtypic binding to NAs from influenza A group 1 (N1, N4, N5, N8), group 2 (N2, N3, N6, N7, N9), and influenza B viruses (Fig. 1A).

Single plasmablasts were sorted from peripheral blood mononuclear cells (PBMCs) on day 5 after the onset of symptomatic illness, and the corresponding immunoglobulin heavy and light variable (IGHV and IGLV) chain genes were cloned and expressed as previously described (*23*). The mAbs were then screened for binding to recombinant H3 HA, N2 NA, nucleoprotein, and matrix protein 1. Three antibodies from this screening (of 45 mAbs isolated)—1G04, 1E01, and 1G01—bound to the N2 NA of the seasonal influenza virus strain A/Hong Kong/4801/2014 (H3N2), which is presumably closely related to the strain that caused the infection. The antigen-binding fragments (Fabs) of 1G04, 1E01, and 1G01 bound with low nanomolar affinities to the N2 NA of A/Hong Kong/4801/2014 (fig. S1). 1G04, 1E01, and 1G01 form a three-member clonal family where IGHV3-20 and IGKV1-9 encode the IGHV and IGLV genes, respectively (Fig. 1B). The plasmablasts producing the three mAbs most likely originated from memory B cells, as evidenced by the accumulation of many somatic hypermutations, particularly in the IGHV genes. Alignment of the heavy chain amino acid sequences of each mAb to the germline gene, IGHV3-20*04, showed that 1G04, 1E01, and 1G01 differ at 12, 14, and 19 positions, respectively (Fig. 1B).

Upon further characterization, the three antibodies displayed broad binding to recombinant N2 NA (group 2) from seasonal and avian influenza viruses (Fig. 1A and table S1). Furthermore, 1G04 showed some cross-reactivity to N3 and N6 (group 2) and N1, N5, and N8 (group 1), as well as weak binding to influenza B NA (Fig. 1C). 1E01 showed a broader binding pattern that included group 1 NAs (N1, N5, and N8) and group 2 NAs (N3, N6, N7, and N9), as well as strong binding to influenza B NA from the B/Victoria/2/87 lineage and weak binding to the NA of the B/Yamagata/16/88 lineage. Finally, mAb 1G01 showed the broadest binding activity that covered all group 1 NAs (N1, N4, N5, and N8) and group 2 NAs (N2, N3, N6, N7, and N9), as well as NAs from both influenza B virus lineages (Fig. 1C).

We next examined the functional capacity of the three mAbs in an enzyme-linked lectin assay (ELLA) that measures neuraminidase inhibition (NI) (Figs. 1D and 2A, fig. S2, and table S2). 1G04 inhibited N2, N3, and certain N1 NAs, whereas 1E01 inhibited group 2 NAs, N1 NAs, and B/Victoria/2/87 lineage NAs. 1G01 inhibited the activity of all A/group 2 and group 1 NAs and B/Victoria/2/87 lineage NAs (Fig. 1D, Fig. 2A, and fig. S2). NA activity can be inhibited by antibodies binding directly to epitopes within the enzyme active site or through steric hindrance when antibodies bind proximal to the active site. However, NI through steric hindrance is only observed in the ELLA when larger substrates, such as fetuin, are used. When small molecules are used as substrate, as in an NA-Star assay, antibodies that do not bind directly to the active site might not inhibit (*9*, *10*). When tested in the NA-Star assay, all three mAbs potently inhibited NA activity, hinting at a binding footprint within the NA active site (Fig. 2B and fig. S3). For 1G01 [50% inhibitory concentration (IC_{50}) of

[1]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [2]Department of Biotechnology, University of Natural Resources and Life Sciences, A-1190 Vienna, Austria. [3]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [4]Division of Immunobiology, Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA. [5]Department of Pediatrics, Columbia Irving Medical Center, New York, NY 10032, USA. [6]NewYork-Presbyterian Morgan Stanley Children's Hospital, New York, NY 10032, USA. [7]Division of Emergency Medicine, Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA. [8]Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: wilson@scripps.edu (I.A.W.); ellebedy@wustl.edu (A.H.E.); florian.krammer@mssm.edu (F.K.)

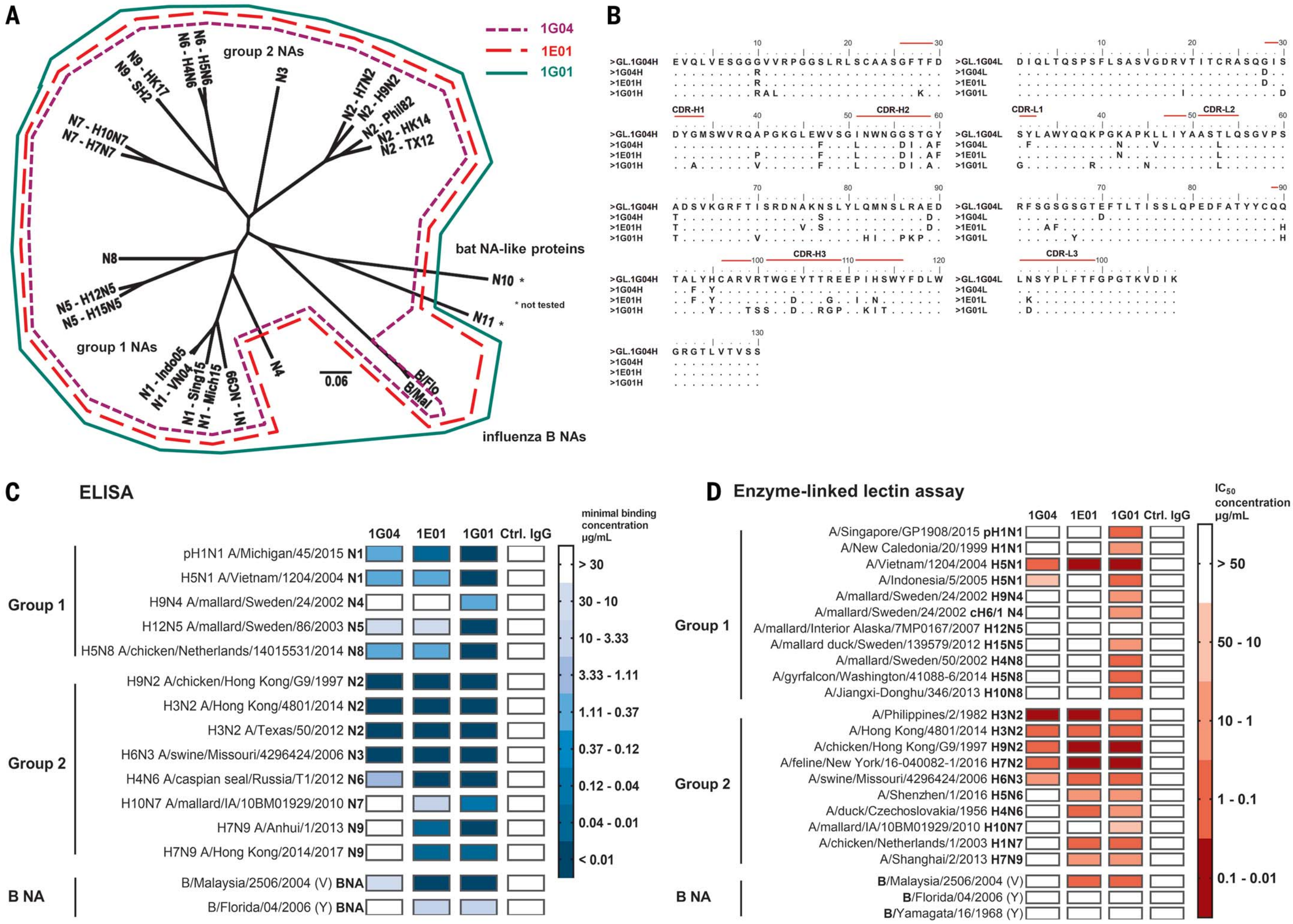


Fig. 1. Breadth of mAbs 1G04, 1E01, and 1G01. (**A**) Phylogenetic tree of influenza A and B virus NAs. The reactivity breadth of the three mAbs is indicated. The scale bar represents a 6% change in amino acids. The tree was built using amino acid sequences in ClustalOmega and visualized in FigTree. (**B**) Alignment of the amino acid sequences of each mAb heavy chain (top) and light chain (bottom) to their closest germline immunoglobulin genes as identified by NCBI IgBlast. (**C**) Heat map of antibody binding to recombinant protein in enzyme-linked immunosorbent assay. The minimal binding concentration is indicated. (**D**) Heat map of antibody activity in ELLA NI assays. The IC_{50} is indicated. For reassortant strains, virus strain names correspond to the NA of the virus used.

1.35 nM], this notion was confirmed using the Fab instead of the full antibody in ELLA to remove any steric hindrance contribution by the fraction crystallizable (Fc) region to inhibition. The 1G01 Fab still displayed potent inhibition (IC_{50} of 7.62 nM), which indicates that the antibody may be directly targeting the active site (Fig. 2C). Interestingly, a germline-reverted version of 1G04 also potently inhibited NA activity in the ELLA assays (Fig. 2D). All three antibodies inhibited an oseltamivir-resistant H3N2 virus with potency similar to what we observed with a susceptible control virus (fig. S4). Typically, anti-NA antibodies do not show activity in neutralization assays in vitro (*9*), but all three mAbs showed strong inhibitory activity in assays against various N2 viruses (Fig. 2E and fig. S5). Although this effect is not due to inhibition of viral entry into cells, it reflects the strong NI activity of the mAbs in blocking virus egress and replication necessary for a positive readout in that assay. Of note, the mAbs neutralized and inhibited the H3N2 A/Philippines/2/1982 virus more potently than it did the more recent H3N2 A/Hong Kong/4801/2014 virus (Fig. 1D, Fig. 2E, and fig. S5), which may be due to the patient's initial exposure/imprinting to H3N2 viruses in the 1980s when the patient was born. In addition to Fab-based antiviral activity, Fc receptor (FcR)–mediated effector functions, such as antibody-dependent cellular cytotoxicity (ADCC), are important for some broadly protective anti-HA antibodies and have also been detected for anti-NA antibodies (*9*, *24*). All three mAbs showed activity against a panel of different NAs in an ADCC bioreporter assay, even when they exhibited weak binding (e.g., 1G04 binding to influenza B NA) (Fig. 2F and fig. S6).

To elucidate the epitopes of the three mAbs and the structural basis of their broad protection, we determined crystal structures of 1G04 with the N9 NA from the recent H7N9 epidemic isolate A/Hunan/02650/2016 (Hunan N9) at 3.45 Å resolution (1G04 has binding affinity to this specific strain of N9), 1E01 in complex with the N2 NA of the 1957 H2N2 pandemic isolate A/Japan/305/1957 (Japan57 N2) at 2.45 Å resolution, and 1G01 in complex with the N1 NA of the 2009 pandemic H1N1 isolate A/California/04/2009 (CA04 N1) at 3.27 Å resolution (Fig. 3 and table S3). In all three Fab-NA complex structures, one Fab bound to one NA protomer of the NA tetramer. All three Fabs bound to the NAs such that they fully blocked the NA active site (Fig. 3). 1G04, 1E01, and 1G01 recognized their epitopes using both heavy and light chains; in all three antibodies, the 21–amino acid complementarity-

determining region (CDR) H3 played an especially dominant role in their NA interactions. The total surface area on the NA buried by 1G04, 1E01, and 1G01 was 1030 Å^2, 900 Å^2, and 800 Å^2, respectively, of which 87%, 84%, and 77% arose from the heavy chain, and 67%, 66%, and 77% from the extended CDR H3 loop.

Antibody 1G04 interacted with Hunan N9 NA using CDRs L1, L2, H1, H2, and H3 (Fig. 3A, fig. S7, and table S4). The CDR H3 of 1G04 contributed most of the Fab interactions to 20 N9 NA residues, including 10 active-site residues, which are conserved in group 1, group 2, and influenza B NAs: Arg^{118}, Glu^{119}, Asp^{151}, Arg^{152}, Ile^{222}, Arg^{224}, Glu^{276}, Glu^{277}, Arg^{371}, and Tyr^{406}. CDRs L1, L2, H1, and H2 contacted some nonconserved N9 NA residues. Similarly, 1E01 interacted with Japan57 N2 using the same five CDR loops as 1G04. The CDR H3 of 1E01 dominated the interaction with 16 N2 NA residues, including all of the conserved active-site epitope residues recognized by 1G04 except for Glu^{276} (Fig. 3B, fig. S8, and table S5). Interestingly, 1G01 bound CA04 N1 NA using only CDRs L1, L2, and H3, as well as one residue, Tyr^{67}, from the framework region (FR) FR3. The 21-residue CDR H3 of 1G01 interacted with 19 CA04 N1 NA residues, which included 13 NA-conserved active-site residues (10 of which are conserved with 1G04) or second-shell residues, including Leu^{134}, Arg^{156}, and Trp^{178} (Fig. 3C, fig. S9, and table S6). The epitopes of these three antibodies are distinct from the structurally characterized mouse NA antibodies, which bind to epitopes distant from the active site or on the rim of the active site (*12*, *25–27*).

Next, we assessed the protective capacity of the three mAbs in vivo in the mouse model. We tested the ability of the mAbs to protect against viruses expressing a human N2, an avian N2, a swine N3, and avian N6, N7, and N9 NAs (all group 2); a human N1 and avian N1, N4, N5, and N8 NAs (all group 1); and an influenza B NA from the B/Victoria/2/87 lineage (Fig. 4, A and C to M, and table S7). All virus strains were lethal in the mouse challenge models except for avian N2 and N6, in which severe weight loss, but no mortality, was observed in the animals that received the control antibody. Animals received mAbs at 5 mg/kg 2 hours before infection, and morbidity (weight loss) and mortality were then monitored for 14 days. The mAbs protected from both weight loss and mortality in a manner that was consistent with their reactivity patterns, which indicates that the ability of the antibody to bind substrate in vitro is predictive of its ability to protect in vivo. In some cases (e.g., for 1G04 and the H4N6 and B/Malaysia/2506/2004 challenges) (Fig. 4, E and M), protection was observed despite the absence of strong NI activity. However, binding can mediate effector function (fig. S6), which can, in combination with residual NI, likely also provide protection from challenge. Notably, 1G01 provided full protection from lethality against all challenge viruses. Except for the N4 challenge (Fig. 4J), where some transient weight loss was observed, 1G01 provided complete protection against weight loss. Furthermore, the protective effect of 1G01 was tested in a dose de-escalation study with an H3N2 challenge in both BALB/c and DBA/2J mice. A dose as low as 0.3 mg/kg was fully protective for BALB/c mice, and 1 mg/kg protected 80% of the DBA/2J (fig. S10). We also monitored virus replication in the lungs of challenged animals on day 3 and day 6 after infection and did not detect any replicating virus using

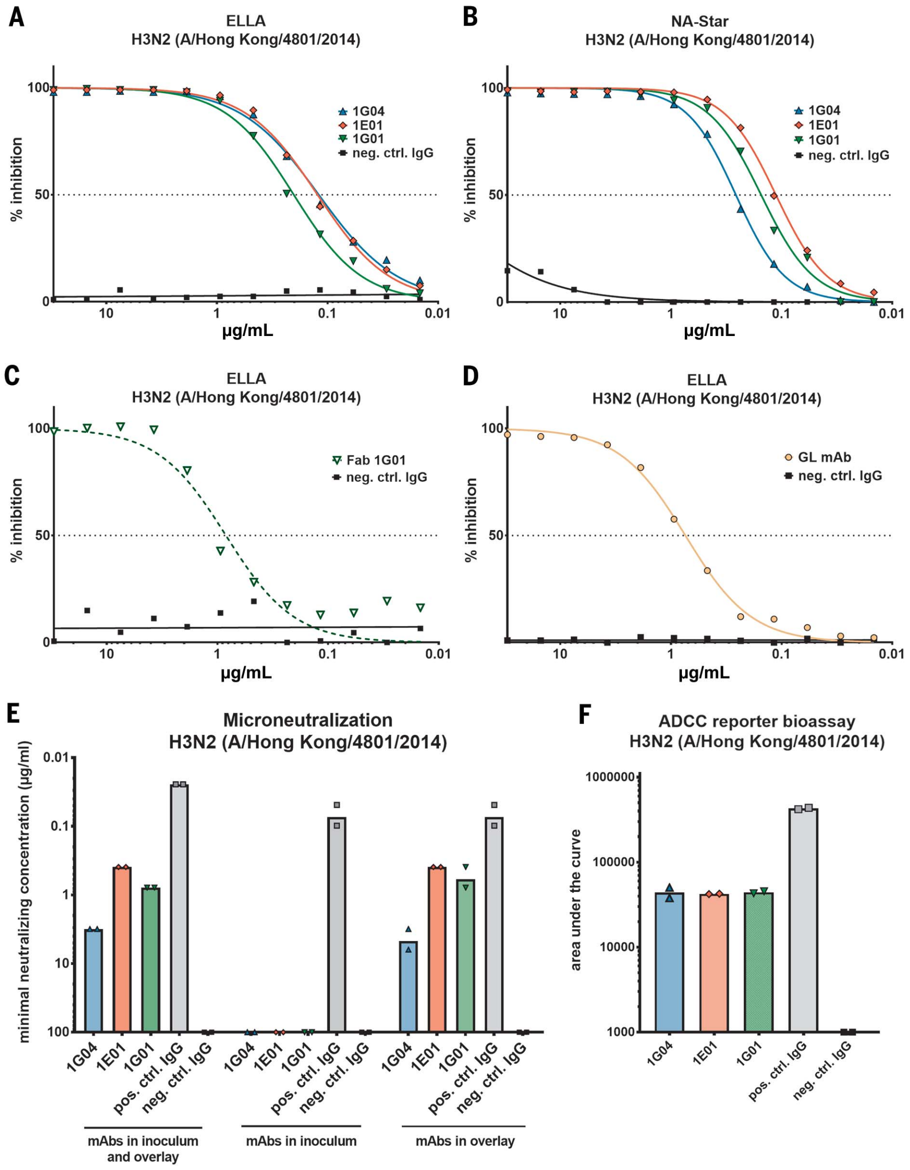


Fig. 2. In vitro functionality of mAbs 1G04, 1E01, and 1G01. (**A**) Representative NA inhibition curves in an ELLA assay (large substrate, steric hindrance–sensitive) against H3N2 strain A/Hong Kong/4801/2014. (**B**) Activity of the same mAbs against A/Hong Kong/4801/2014 in an NA-Star assay (small substrate, steric hindrance–insensitive). (**C**) Activity of the Fab of mAb 1G01 in an ELLA assay against A/Hong Kong/4801/2014. (**D**) A germline (GL) version of 1G04, representing all three antibodies, is also active against H3N2 strain A/Hong Kong/4801/2014. (**E**) All three mAbs are active in a microneutralization assay against H3N2 strain A/Hong Kong/4801/2014 when antibody is added to the inoculum and overlay. The NA antibodies cannot neutralize the virus when added to the inoculum only, but antibodies prevent viral egress when added to the overlay. The minimal neutralizing concentration is defined as the lowest antibody concentration at which no hemagglutination is detected. (**F**) ADCC reporter bioassay activity of the three mAbs. MAb CR9114, which has known ADCC reporter activity against H3N2, was used as positive control.

a plaque assay; this finding suggests that the strong in vitro inhibiting and neutralizing activity of the three mAbs translates to sterilizing immunity in vivo, at least on the sampled days (Fig. 4B). This observation was further corroborated by measuring titers using a limiting dilution method in embryonated eggs, which has lower sensitivity. Only residual titers were found on day 3 after infection and no virus was detected in any treated mouse on day 6 after infection (fig. S11). To determine whether the mAbs had therapeutic potential, we infected mice with a lethal dose of H3N2 virus and treated them with antibody 48 or 72 hours after infection (Fig. 4, N and O). Although transient weight loss was observed, all animals recovered when treated with a low mAb dose of 5 mg/kg.

The isolation of these three mAbs illustrates that humans can make highly potent heterosubtypic anti-NA antibodies that can inhibit influenza A and B viruses. The three mAbs represent 6.67% of the generated mAbs from this patient, indicating that, at least in this individual, they are relatively common. However, earlier antibody characterization and vaccination studies indicated that although protection can be very broad within a subtype, immunity to NA is not heterosubtypic (compared to anti-HA stalk antibodies that typically bind across different subtypes) (*9*–*15*). Attempts to raise heterosubtypic NA mAbs in rabbits against a conserved, linear peptide yielded mAbs with low NI activity and little protective effect (*28*, *29*). A few human NA mAbs have recently been reported that show some cross-reactive binding, but they do not have any antiviral activity in vitro or in vivo (*10*). Of note, there is some serological evidence with polyclonal serum that natural infection in humans can induce broadly reactive anti-NA antibodies (*30*, *31*). The mAbs we report here are highly potent and mediate strong protection in vivo, with 1G01, which has the highest number of somatic hypermutations and the smallest footprint, exhibiting the broadest binding.

On the basis of the reactivity profiles, it seems likely that the B cell clones encoding these mAbs were initially engaged by an H3N2 infection and have acquired affinity to N1 and influenza B virus NA antigens through subsequent exposures of this individual to H1N1 and influenza B viruses. It is conceivable that such antigenically broad exposures allowed these mAbs to recognize other influenza A virus NA subtypes to which the individual was never exposed. The recent H3N2 infection of this individual then recalled these clones. In fact, serum from the patient taken at 5 days after symptom onset displayed cross-reactivity against various NAs and also exhibited high hemagglutination and microneutralization titers against the H3N2 virus A/Hong Kong/4801/2014 (320 and 2560, respectively). An NI titer of 536 (against an H6N2 virus, A/Texas/50/2012) was also observed (fig. S12). It is possible that the pre-exposure serum concentrations of these antibody species in this individual were not high enough to mediate protection. In addition, the patient presented with several underlying risk factors such as obesity, diabetes, hypertrophic cardiomyopathy, and asthma, which could explain why the person became infected.

Neuraminidase is a validated drug target, and several small molecules that inhibit its activity are licensed as influenza therapeutics (*32*). In this work, several attempts to generate H3N2 escape mutants were unsuccessful. Like the three mAbs reported here, these small molecules target the active site of the NA.

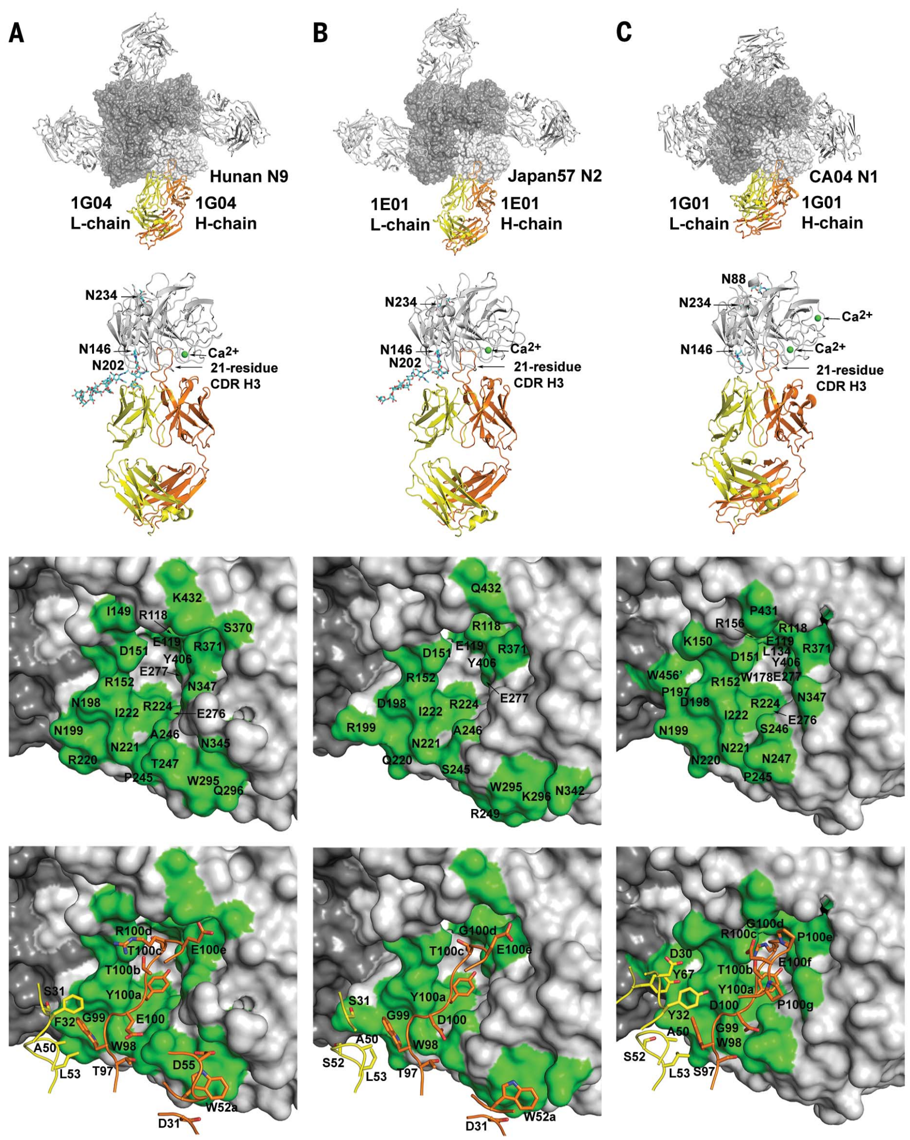


Fig. 3. Crystal structures of 1G04, 1E01, and 1G01 Fabs in complexes with NAs. (**A**) Crystal structure of 1G04 with Hunan N9 NA at 3.45 Å resolution. From top to bottom: the NA tetramer with one Fab bound to each protomer of the NA tetramer, the NA protomer with one Fab, the epitope on the NA, and the Fab paratope on the NA. (**B**) Crystal structure of 1E01 with Japan57 N2 NA at 2.45 Å resolution. From top to bottom, as for (A). (**C**) Crystal structure of 1G01 with CA04 N1 NA at 3.27 Å resolution. From top to bottom, as for (A). For all panels, one NA-Fab protomer is colored with NA in light gray, Fab light chain (L-chain) in yellow, and Fab heavy chain (H-chain) in orange. The NA and Fab in the other protomers are in dark and light gray, respectively. N-linked glycans are shown in stick representation with cyan carbon atoms. Calcium ions are shown as green spheres. The molecular surface depicting the epitope is colored in green. Amino acid abbreviations: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; W, Trp; Y, Tyr.

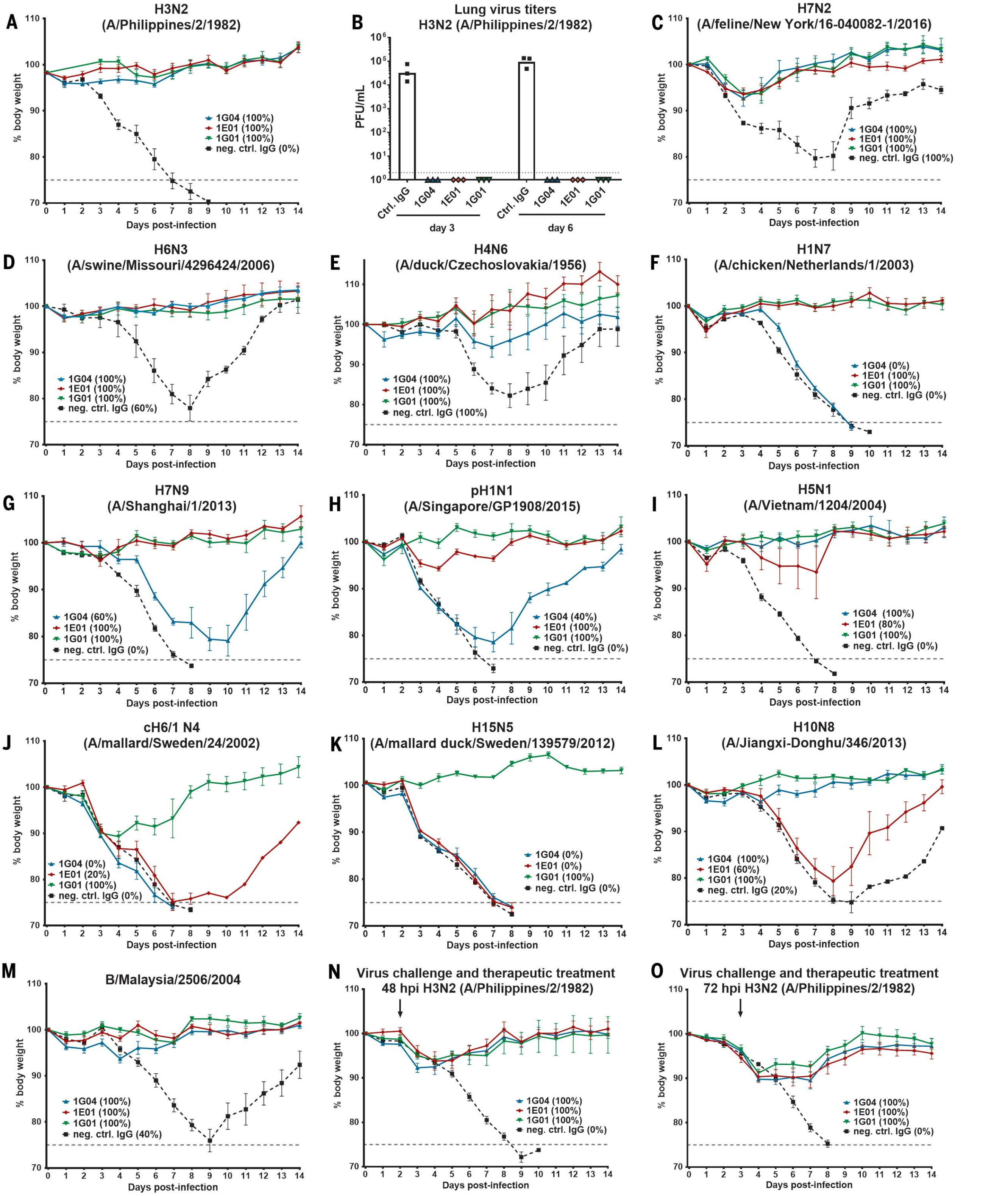


Fig. 4. In vivo protection by mAbs 1G04, 1E01, and 1G01. (**A** and **C** to **M**) Efficacy of the three mAbs in a mouse model against challenge with viruses expressing group 1, group 2, and influenza B virus NAs. Animals were treated with mAbs (5 mg/kg) intraperitoneally 2 hours before intranasal virus challenge. Five animals per group were used. (**B**) Lung titers of animals treated prophylactically with mAbs [as described for (A)] on day 3 and day 6 after infection. Three mice per group were used. (**N** and **O**) We also tested therapeutic treatment of 5 mg/kg at 48 hours (N) and 72 hours (O) after infection. The treatment time points are indicated with errors. Five animals were used per group. For (A) and (C) to (O), percent weight loss is shown and percent survival is indicated in the respective panels. Weight loss was monitored daily except for the 1G01 group in (A) on days 1 and 2 due to technical malfunctioning of a cage. For reassortant strains, virus strain names correspond to the NA of the virus used.

Therefore, and because of the extensive breadth of these mAbs, they could potentially be used as antivirals for treatment of seasonal, pandemic, and zoonotic influenza virus infection in humans. In addition to the potential of these mAbs as therapeutics, they might also be useful for antibody-guided vaccine design. Currently licensed influenza virus vaccines, including inactivated and live attenuated vaccines, are poor inducers of anti-NA immunity (*8*). However, recombinant NA-based vaccines, as well as NA virus-like particles, have been shown to induce high titers of anti-NA antibodies that can broadly protect in lethal mouse and ferret challenge models (*13*–*15*). Furthermore, recombinant NA-based vaccines have been shown to inhibit influenza virus transmission in the guinea pig model (*16*). Broad protection within a subtype was observed in these studies, but heterosubtypic immunity was not detected, although it was explicitly tested in the mouse model (*13*). However, sequential prime-boost regimens with recombinant NA vaccines that include different subtypes or scaffolded/stabilized constructs that represent the 1G01 epitope might result in the induction of 1G01-like antibodies.

All three mAbs that we isolated show broad binding, with 1G01 inhibiting all influenza A virus NA subtypes and influenza B virus NAs, making these mAbs extremely promising candidates for therapeutic development. These antibodies are potent inhibitors of NA activity in vitro and provide broad protection from mortality and morbidity in vivo. The discovery of these mAbs raises the hope that similar antibodies can be induced in the population if the right vaccination regimen is administered. Knowledge of the binding mode and epitope of these mAbs may then guide the development of NA-based universal influenza virus vaccines.

REFERENCES AND NOTES

1. F. Krammer, P. Palese, *Nat. Rev. Drug Discov.* **14**, 167–182 (2015).
2. E. J. Erbelding *et al.*, *J. Infect. Dis.* **218**, 347–354 (2018).
3. F. Krammer, *Nat. Rev. Immunol.* **19**, 383–397 (2019).
4. D. C. Ekiert *et al.*, *Science* **333**, 843–850 (2011).
5. D. C. Ekiert *et al.*, *Science* **324**, 246–251 (2009).
6. C. Dreyfus *et al.*, *Science* **337**, 1343–1348 (2012).
7. D. Corti *et al.*, *Science* **333**, 850–856 (2011).
8. F. Krammer *et al.*, *mBio* **9**, e02332-17 (2018).
9. T. J. Wohlbold *et al.*, *Nat. Microbiol.* **2**, 1415–1424 (2017).
10. Y. Q. Chen *et al.*, *Cell* **173**, 417–429.e10 (2018).
11. L. Jiang *et al.*, *J. Virol.* **90**, 117–128 (2015).
12. H. Wan *et al.*, *Nat. Commun.* **6**, 6114 (2015).
13. T. J. Wohlbold *et al.*, *mBio* **6**, e02556 (2015).
14. G. E. Smith *et al.*, *Virology* **509**, 90–97 (2017).
15. W. C. Liu, C. Y. Lin, Y. T. Tsou, J. T. Jan, S. C. Wu, *J. Virol.* **89**, 7224–7234 (2015).
16. M. McMahon *et al.*, *mBio* **7**, 00560-19 (2019).
17. A. S. Monto *et al.*, *J. Infect. Dis.* **212**, 1191–1199 (2015).
18. R. B. Couch *et al.*, *J. Infect. Dis.* **207**, 974–981 (2013).
19. M. J. Memoli *et al.*, *mBio* **7**, e00417-16 (2016).
20. N. S. Heaton, D. Sachs, C. J. Chen, R. Hai, P. Palese, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20248–20253 (2013).
21. M. R. Sandbulte *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20748–20753 (2011).
22. Y. Abed, I. Hardy, Y. Li, G. Boivin, *J. Med. Virol.* **67**, 589–595 (2002).
23. J. Wrammert *et al.*, *Nature* **453**, 667–671 (2008).
24. D. J. DiLillo, G. S. Tan, P. Palese, J. V. Ravetch, *Nat. Med.* **20**, 143–151 (2014).
25. W. R. Tulip, J. N. Varghese, W. G. Laver, R. G. Webster, P. M. Colman, *J. Mol. Biol.* **227**, 122–148 (1992).
26. R. L. Malby *et al.*, *Structure* **2**, 733–746 (1994).
27. L. Venkatramani *et al.*, *J. Mol. Biol.* **356**, 651–663 (2006).
28. T. M. Doyle *et al.*, *Biochem. Biophys. Res. Commun.* **441**, 226–229 (2013).
29. T. M. Doyle *et al.*, *Antiviral Res.* **100**, 567–574 (2013).
30. L. Liu *et al.*, *J. Infect. Dis.* **215**, 518–528 (2017).
31. R. Nachbagauer *et al.*, *Nat. Immunol.* **18**, 464–473 (2017).
32. A. C. Hurt, *Trop. Med. Infect. Dis.* **4**, 67 (2019).

ACKNOWLEDGMENTS

We thank A. Javier, M. Amor, and F. Amanat for technical support, and H. Tien for automated crystal screening; the Clinical Microbiology Laboratory, Department of Pathology at the Icahn School of Medicine (M. Gitman, M. Sordillo), the Personalized Virology Initiative (J. Tan, E. Hirsch, J. Polanco, V. Simon), and the Pathogen Surveillance Program (D. Altman, H. van Bakel) for providing two recent influenza B virus isolates, and M. Eichelberger at the FDA for providing an H6N2 reassortant virus [which were generated by reverse genetics plasmids provided by R. Webby (St. Jude Children's Hospital)]; and S. House and the clinical sample collection team in the Washington University Emergency Care and Research Core (ECRC). The EDFLU study was reviewed and approved by the Washington University Institutional Review Board (approval no. 2017-10-220). **Funding:** The Krammer laboratory was supported by NIAID CEIRS contract HHSN272201400008C and NIAID grant R01 AI117287; the Ellebedy laboratory was supported by NIAID R21 AI139813, U01 AI141990, and NIAID CEIRS contract HHSN272201400006C; and the Wilson laboratory was supported in part by NIH R56 AI117675 (to I.A.W.). Human sample collection was supported by a grant from the Washington University Institute of Clinical and Translational Sciences to P.A.M., which is, in part, supported by the NIH/National Center for Advancing Translational Sciences (NCATS), CTSA grant UL1TR002345. X-ray diffraction data were collected at the Advanced Photon Source (APS) beamline 23ID-B and at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2. Use of the APS was supported by the U.S. Department of Energy (DOE), Basic Energy Sciences, Office of Science, under contract no. DE-AC02-06CH11357. GM/CA CAT is funded in whole or in part with federal funds from the National Cancer Institute (Y1-CO-1020) and the NIGMS (Y1- GM-1104). Use of the SSRL, SLAC National Accelerator Laboratory, is supported by the U.S. DOE, Office of Science, Office of Basic Energy Sciences under contract DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institute of General Medical Sciences (including P41 GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS, NIAID, or NIH. **Author contributions:** D.S., X.Z., M.M., T.J.W., R.N., S.S., and W.Y. characterized the antibodies; P.A.M., A.J.S., and J.S.T. isolated the antibodies; and D.S., X.Z., I.A.W., A.H.E., and F.K. conceptualized the study and wrote the manuscript. **Competing interests:** D.S., X.Z., M.M., R.N., J.S.T., A.J.S., P.A.M., T.J.W., S.S., W.Y., F.K., and I.A.W. declare no conflict of interest. A.H.E. is a consultant for Inbios. Washington University has filed a patent application based on the findings reported here with A.H.E., F.K., and I.A.W. as inventors. **Data and materials availability:** Coordinates and structure factors have been deposited in the Protein Data Bank as entries 6Q1Z (1G04 in complex with N9 NA), 6Q20 (1E01 Fab in complex with N2 NA), and 6Q23 (1G01 Fab in complex with N1 NA). Antibody sequences are deposited on GenBank [accession nos. MN013068 (1G01 VH), MN013072 (1G01 Vκ), MN013070 (1E01 VH), MN013070 (1E01 Vκ), MN013069 (1G04 VH), and MN013073 (1G04 Vκ) available from GenBank/EMBL/DDBJ]. All data are available in the main text or the supplementary materials. mAbs 1G01, 1G04, and 1E04 used in the analysis are available from the participating laboratories under standard academic material transfer agreements.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/499/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S7
References (*33*–*45*)

View/request a protocol for this paper from *Bio-protocol*.

16 May 2019; accepted 27 September 2019
10.1126/science.aay0678

STRUCTURAL BIOLOGY

Cryo-EM structures of the human cation-chloride cotransporter KCC1

Si Liu[1,2]*, Shenghai Chang[1,3]*, Binming Han[4]*, Lingyi Xu[1,5], Mingfeng Zhang[6], Cheng Zhao[1,6], Wei Yang[6], Feng Wang[7], Jingyuan Li[4]†, Eric Delpire[8]†, Sheng Ye[2,5]†, Xiao-chen Bai[9]†, Jiangtao Guo[1,6]†

Cation-chloride cotransporters (CCCs) mediate the coupled, electroneutral symport of cations with chloride across the plasma membrane and are vital for cell volume regulation, salt reabsorption in the kidney, and γ-aminobutyric acid (GABA)–mediated modulation in neurons. Here we present cryo–electron microscopy (cryo-EM) structures of human potassium-chloride cotransporter KCC1 in potassium chloride or sodium chloride at 2.9- to 3.5-angstrom resolution. KCC1 exists as a dimer, with both extracellular and transmembrane domains involved in dimerization. The structural and functional analyses, along with computational studies, reveal one potassium site and two chloride sites in KCC1, which are all required for the ion transport activity. KCC1 adopts an inward-facing conformation, with the extracellular gate occluded. The KCC1 structures allow us to model a potential ion transport mechanism in KCCs and provide a blueprint for drug design.

The human solute carrier 12 (*SLC12*) gene family encodes the cation-chloride cotransporters (CCCs) that mediate the coupled, electroneutral symport of Na^+ and/or K^+ with Cl^- across the plasma membrane (*1*, *2*). Defined by their transport properties and amino acid sequences, CCCs can be divided into several branches, including two Na-K-2Cl cotransporters (NKCC1 and NKCC2), one Na-Cl cotransporter (NCC), and four K-Cl cotransporters (KCC1 to KCC4) (fig. S1). CCCs are expressed in many tissues and organs, especially in kidney and the nervous system, and play important roles in a variety of physiological processes. NKCCs and KCCs participate in cell volume regulation by transporting ions across the plasma membrane to reduce the osmotic difference when cells are exposed to hypertonic or hypotonic environments (*3*). In kidney, NCC and NKCC2 are major carriers for the reabsorption of NaCl in the distal convoluted tubule and the thick ascending loop of Henle, respectively (*4*). In central neurons, NKCC1 and KCC2 collaboratively set the intracellular Cl^- concentration and thereby regulate γ-aminobutyric acid (GABA)– and glycine-mediated synaptic inhibition (*5*).

Considering their important functions in ion transport, it is not surprising that mutations in CCCs lead to a wide spectrum of human diseases (*6*–*10*). For example, five single amino acid substitutions in KCC2 can cause human epilepsy. R952H (Arg^{952}→His) and R1049C (Arg^{1049}→Cys) mutations give rise to idiopathic generalized epilepsy (*9*), whereas L311H (Leu^{311}→His), L426P (Leu^{426}→Pro), and G551D (Gly^{551}→Asp) substitutions induce epilepsy of infancy with migrating focal seizures (*10*). Therefore, CCCs are important drug targets. The diuretics furosemide and thiazide are used as antihypertensive drugs by inhibition of NKCC2 and NCC, respectively (*11*). Recently, efforts have been made to develop KCC2-specific agonists and NKCC1-specific antagonists to treat epilepsy diseases (*12*, *13*). Drugs targeting KCC2 or NKCC1 may also be effective in treating other diseases by regulating GABAergic inhibition. For example, a KCC2 agonist, CLP290, promoted functional recovery after spinal cord injury (*14*).

Although the physiology, biochemistry, and pharmacology of different CCC members have been extensively studied, our knowledge of their ion cotransport mechanism is hindered by the lack of high-resolution structures. Structures are available for the *Methanosarcina acetivorans* CCC C-terminal domain (*15*) and *Danio rerio* NKCC1 (DrNKCC1) (*16*). Although the DrNKCC1 structure reveals the architecture and ion binding sites of the sodium-driven CCCs, the structures and ion cotransport mechanisms of the potassium-driven KCCs remain unknown. Here, we present structural studies of human KCC1.

We determined three cryo–electron microscopy (cryo-EM) structures of human KCC1: one in 150 mM KCl and detergent glyco-diosgenin (GDN) at 2.9-Å resolution, one reconstituted into lipid nanodisc in 150 mM KCl at 2.9-Å resolution, and one in 150 mM NaCl and GDN at 3.5-Å resolution (figs. S2 to S7 and table S1). Because the overall architectures of the three structures are very similar (fig. S8), we mainly focus on the structure in KCl and GDN in this report. The cryo-EM map in the transmembrane domain (TMD) and extracellular domain (ECD) is of high quality, enabling us to build an accurate model containing residues 116 to 652 (Fig. 1, A and B, and fig. S3). Owing to the flexibility of the linker that connects the TMD and C-terminal domain (CTD), the CTD is resolved at a lower resolution and is not modeled.

The TMD of KCC1 contains 12 transmembrane helices (TMs), with both the amino and carboxyl termini on the intracellular side (Fig. 1, C and D). TM1 to TM5 and TM6 to TM10 are inverted structural repeats related by a pseudo–twofold symmetry axis parallel to the membrane plane, with a root-mean-square deviation of 4.6 Å over 136 Cα atoms (fig. S9). Whereas TM1, TM2, TM6, and TM7 form the "core domain," TM3 to TM5 and TM8 to TM10 form the "scaffold domain," providing ion transport cavities between the core domain and the scaffold domain. TM11 and TM12 wrap around TM10 at one side of the transporter (Fig. 1D). The KCC1 TMD adopts a LeuT-like fold, which is shared among a superfamily of secondary transport proteins termed amino acid, polyamine, and organocation (APC) transporters (*17*).

The ECD, formed by a 120-residue-long linker between TM5 and TM6, is a characteristic and conserved feature of KCCs, whereas NKCCs and NCC have much shorter linkers (fig. S1). The ECD consists of two pairs of antiparallel β strands (β1-β4 and β2-β3) and four short α helices (EL2 to EL5). The β2-β3 hairpin extends like an antenna from the ECD (Fig. 1D). The relatively rigid architecture of the ECD is stabilized by two disulfide bonds, C308–C323 and C343–C353 (Fig. 1D). Mutations of each of the four highly conserved cysteine residues abolish KCC2 transport activity, suggesting that the ECD folded structure is important for function (*18*). Two asparagine residues, Asn^{312} and Asn^{361}, are glycosylated with visible density for the covalently linked *N*-acetylglucosamine moiety of the sugar (Fig. 1B and fig. S3B).

The cryo-EM structure of KCC1 reveals that both the TMD and ECD participate in dimeric assembly. At the dimer interface of the ECD, there are a pair of hydrogen bonds between the carbonyl oxygen of Ser^{403} and the amine group of Lys^{405} from neighboring subunits

[1]Department of Biophysics, Department of Pathology of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China. [2]Tianjin Key Laboratory of Function and Application of Biological Macromolecular Structures, School of Life Sciences, Tianjin University, 92 Weijin Road, Nankai District, Tianjin 300072, China. [3]Center of Cryo-Electron Microscopy, Zhejiang University School of Medicine, Hangzhou 310058, China. [4]Zhejiang Province Key Laboratory of Quantum Technology and Device, Institute of Quantitative Biology, Department of Physics, Zhejiang University, Hangzhou, Zhejiang 310027, China. [5]Life Sciences Institute and Innovation Center for Cell Signaling Network, Zhejiang University, Hangzhou, Zhejiang 310058, China. [6]Department of Biophysics, Institute of Neuroscience, NHC and CAMS Key Laboratory of Medical Neurobiology, Zhejiang University School of Medicine, Hangzhou 310058, China. [7]Wuxi Biortus Biosciences Co. Ltd., 6 Dongsheng West Road, Jiangyin 214437, China. [8]Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, TN 37232, USA. [9]Departments of Biophysics and Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: jiangtaoguo@zju.edu.cn (J.G.); xiaochen.bai@utsouthwestern.edu (X.B.); sye@tju.edu.cn (S.Y.); eric.delpire@vanderbilt.edu (E.D.); jingyuanli@zju.edu.cn (J.L.)

(fig. S10A). In addition, hydrophobic interactions between the side chains of Pro[355], Pro[402], and Leu[404] from both ECDs strengthen this dimerization.

In the TMD of KCC1, the dimeric interaction involves TM12 and TM9 from the adjacent molecules in a dimer. On the periphery, hydrophobic amino acids from TM12 of one KCC1 molecule interdigitate with nonpolar residues from TM9 of the adjacent molecule, sealing the dimer interface at the edges but leaving an interior hydrophobic cavity between the monomers (fig. S10B). Within the cavity, two tube-like cryo-EM densities with the shape and size of two GDN molecules were observed (fig. S3C). The tail of the putative GDN inserts in the hydrophobic cavity and mediates hydrophobic interactions (fig. S10B). In comparison with other APC transporters such as LeuT (*17*) and the *Vibrio parahaemolyticus* sodium-galactose symporter (vSGLT) (*19*), KCC1 has relatively loose dimeric packing in the TMD (fig. S11).

Multiple biochemical studies have shown that CCCs form dimers in vivo (*20*–*24*). Consistently, three lines of evidence support that this dimeric assembly of KCC1 involving the ECD and the TMD is stable and therefore probably physiologically relevant. First, this dimeric state is detergent-independent, because KCC1 has similar retention volumes in gel filtration in various detergents such as *n*-dodecyl-β-D-maltopyranoside (DDM), *n*-decyl-β-D-maltopyranoside (DM), GDN, and lauryl maltose neopentyl glycol (LMNG) (fig. S12A). Second, single or multiple mutations at the dimer interface in the TMD were unable to disrupt the dimeric assembly, confirming the strong dimeric interactions (fig. S12, B and C).

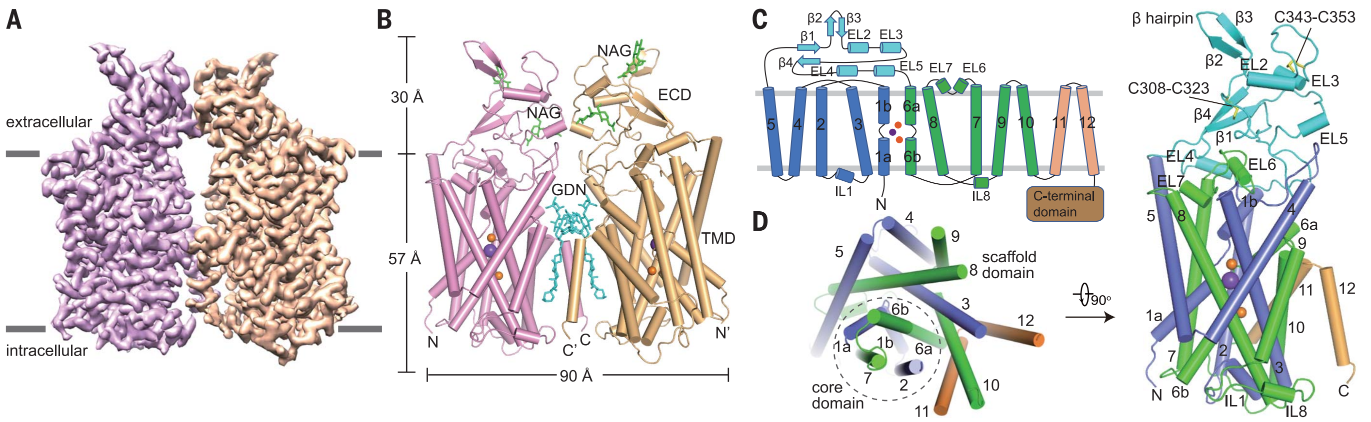
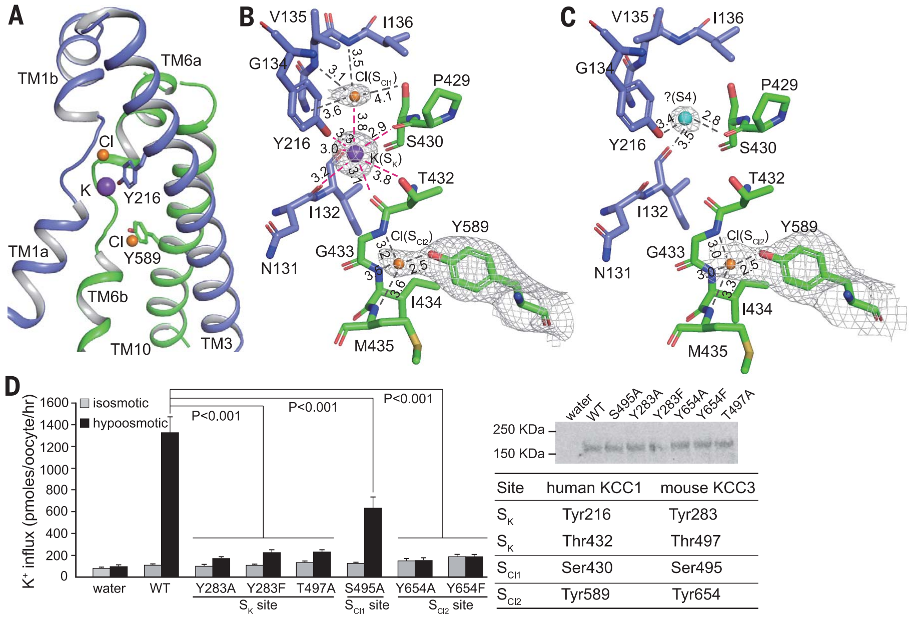


Fig. 1. Overall structure of human KCC1 determined in KCl and GDN. (**A**) Side view of a three-dimensional reconstruction of KCC1 with each subunit individually colored. The gray bars on either side of the structure define the position (top and bottom) of the cell membrane. (**B**) Cartoon representation of the KCC1 dimer in the same orientation as the electron microscopy maps in (A). *N*-Acetylglucosamine (NAG) and putative GDN molecules are shown as green and cyan sticks, respectively. The three ions are shown as purple (K^+) and orange (Cl^-) spheres. N and N′ indicate the N termini; C and C′ indicate the C termini. (**C**) Topology and domain arrangement of the KCC1 subunit. The three ions are shown as purple (K^+) and orange (Cl^-) circles. (**D**) Structure of one KCC1 subunit in intracellular (left) and side (right) views. Each domain is colored the same as in (C). Side chains of two pairs of Cys residues that form disulfide bonds are shown as yellow sticks. Numbers indicate transmembrane helices. TM1, TM2, TM6, and TM7 form the core domain; TM3 to TM5 and TM8 to TM10 form the scaffold domain.

Fig. 2. The ion binding sites in KCC1. (**A**) The potassium and chloride ions bind at the breakage points of TM1 and TM6 of KCC1 in KCl and GDN. Two tyrosine residues involved in the ion coordination are shown as sticks. (**B** and **C**) Magnified views of the potassium and chloride binding sites in the KCC1 structure in KCl (B) or NaCl (C) from the same orientation. In both (B) and (C), all electron densities are shown at the same contour level of 4.5σ. Numbers show distances in angstroms between the ions and coordinating atoms. (**D**) K^+ influx in *Xenopus laevis* oocytes injected with wild-type (WT) and mutant mouse KCC3 complementary RNA and measured under isosmotic (basal) and hypoosmotic (stimulated) conditions. The Western blot shows that all mutants are expressed at similar levels as the wild type. The table shows equivalent residues in human KCC1 and mouse KCC3. For more sequence alignments, see fig. S1. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.



Site	human KCC1	mouse KCC3
S_K	Tyr216	Tyr283
S_K	Thr432	Thr497
S_{Cl1}	Ser430	Ser495
S_{Cl2}	Tyr589	Tyr654

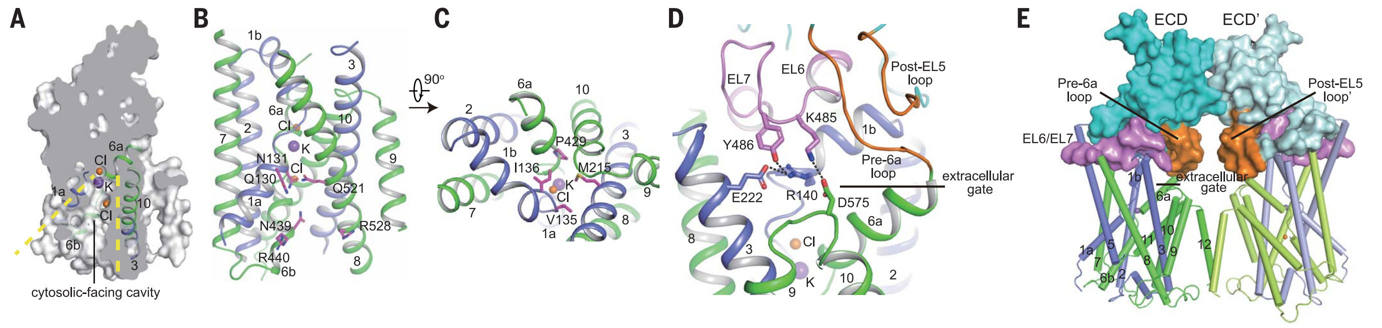


Fig. 3. An inward-facing conformation of KCC1 in KCl and GDN. (**A**) The surface-rendered KCC1 model shows a cytosolic-facing cavity. The yellow dashed lines indicate the opening of the intracellular gate. (**B**) Polar residues on the surface of the cytosolic-facing cavity. (**C**) Hydrophobic residues form a constriction above ion binding sites. (**D**) The extracellular gate is occluded by extracellular linker EL6-EL7 (pink) and pre-6a loop and post-EL5 loop (orange). (**E**) ECD dimerization stabilizes the closed state of the extracellular gates. For clarity, TM4, TM5, TM11, and TM12 are omitted in (B) and (C), as is TM4 in (E).

Third, the cryo-EM structure of KCC1 reconstituted in lipid nanodisc has a very similar overall architecture to that in GDN (figs. S8 and S12D).

In the TMD core domain, TM1 and TM6 are disrupted in the middle by short Gly-containing loops, providing potential ion binding sites approximately halfway across the membrane. The two 2.9-Å resolution structures of KCC1 in KCl reveal three ion binding sites, designated as S_K, S_{Cl1}, and S_{Cl2} (Fig. 2, A and B, and fig. S13). S_K is surrounded by four main-chain carbonyls from both Gly-containing loops (Asn^{131}, Ile^{132}, Thr^{432}, and Pro^{429}) and two hydroxyl groups of Tyr^{216} and Thr^{432}, with coordination distances of around 2.9 to 3.8 Å (Fig. 2B). Because such a local chemical environment clearly favors a cation, we modeled a K^+ at this site, the only cation used in the buffer. The S_{Cl1} site is 3.8 Å above S_K; the ion at this site is mainly coordinated by three main-chain amide groups (Gly^{134}, Val^{135}, and Ile^{136}) and the K^+ at the S_K site. The hydroxyl group of Ser^{430} is about 4 Å from the ion at the S_{Cl1} site, potentially allowing weak coordination. The ion at the S_{Cl2} site is also coordinated mainly by three main-chain amide groups (Gly^{433}, Ile^{434}, and Met^{435}) as well as the hydroxyl group of Tyr^{589} from TM10 (Fig. 2B). Because the chemical environments of S_{Cl1} and S_{Cl2} are suitable for anions, we modeled one Cl^- at each site. A similar Cl^- binding site has been observed in the structure of the prokaryotic Cl^-/H^+ transporter StClC (fig. S14) (*25*).

To test the functions of the ion binding sites, we measured the ion transport activity using a standard K^+ (^{86}Rb) influx assay in *Xenopus laevis* oocytes. Although human KCC1 has low activity (fig. S15A), mutations of the K^+ coordination residue Tyr^{216} or the Cl^- coordination residue Ser^{430} at the S_{Cl1} site did reduce KCC1 activity (fig. S15B). To further confirm the mutants' activity, we studied mouse KCC3, which is a close relative to human KCC1 (78% sequence identity) but with much higher K-Cl cotransport activity (fig. S15C). Mutations of either of the K^+ coordination residues Tyr^{283} (Tyr^{216} in KCC1) or Thr^{497} (Thr^{432} in KCC1), or the Cl^- coordination residue Tyr^{654} (Tyr^{589} in KCC1) at the S_{Cl2} site, abolished the ion transport activities of KCC3 (Fig. 2D). Mutation of the weak Cl^- coordination residue Ser^{495} (Ser^{430} in KCC1) at the S_{Cl1} site also substantially reduced KCC3 activity. This suggests that all three sites are required for K-Cl cotransport activity.

KCCs cotransport K-Cl at a 1:1 ratio (*26*). This raises the questions of why there are two Cl^- sites and which one is the transport site. Three structural features suggest that S_{Cl1} is likely a transport site. First, S_{Cl1} sits deeply in the cytosolic-facing cavity, whereas S_{Cl2} is more exposed to cytosolic solvent. The S_{Cl1} and S_K sites align well with substrates from other APC transporters such as LeuT and GkApcT, a prokaryotic homolog of the SLC7 transporter (fig. S16). Second, the 3.8-Å distance between S_{Cl1} and S_K allows the two ions to coordinate each other and thus to achieve stoichiometric cotransport; by contrast, S_{Cl2} is 7.8 Å below S_K, too far to allow direct interaction with the K^+. Third, Cl^- binding at S_{Cl2} is independent of K^+ binding at S_K, whereas Cl^- binding at S_{Cl1} is coupled to K^+ binding at S_K, as revealed by the KCC1 structure determined in the presence of NaCl. In the KCC1 structure in NaCl, S_{Cl2} is occupied by Cl^-, but no Cl^- binds at the S_{Cl1} site; instead, a fourth site (S4) was observed between S_K and S_{Cl1} (Fig. 2C and supplementary text). Binding of Cl^- at the S_{Cl2} site stabilizes the architecture of the K^+ binding site by direct coordination with the main-chain amide group of Gly^{433} (Fig. 2B). In addition, it strengthens interactions between TM6a, TM6b, and TM10 and thus could couple their conformational changes during the ion transport process. Tyr^{589} has a coordination distance of 2.5 Å to Cl^- at S_{Cl2}; mutation of Tyr^{589} disrupts the S_{Cl2} site, which would likely affect the structural integrity of the K^+ binding site or the coupling of TM6a, TM6b, and TM10 during the ion transport process, resulting in loss of activity.

To confirm the three ion binding sites and to probe the function of the S_{Cl2} site, we performed molecular dynamic simulations (fig. S17). In an equilibrium state without an electrochemical gradient, one K^+ and two Cl^- ions remained stably bound at the S_K, S_{Cl1}, and S_{Cl2} sites throughout a 200-ns simulation (fig. S17A). By contrast, water molecules bound at the S_{Cl1} and S_{Cl2} sites are easy to dissociate (fig. S17, B and C). In the absence of Cl^- binding at the S_{Cl2} site, the glycine-containing loop between TM6a and TM6b relaxes (fig. S17, D and E), resulting in the disturbance of K^+ at the S_K site. The movement of K^+ further affects the binding of Cl^- at S_{Cl1} and even drags it away from its original binding site (fig. S17, D and E). Thus, the removal of Cl^- at S_{Cl2} triggers the correlated dissociation of K^+ from S_K and Cl^- from S_{Cl1} with a correlation coefficient of 0.93. In addition, we performed steered molecular dynamics to pull the K^+ ion from the S_K site. The corresponding potential of mean force profile was used to estimate the binding free energy of K^+, which reduces from -10.28 ± 1.00 kcal/mol to -3.34 ± 1.22 kcal/mol with the removal of Cl^- at the S_{Cl2} site. Thus, Cl^- at S_{Cl2} can facilitate K^+ and Cl^- binding at S_K and S_{Cl1}, respectively.

The KCC1 structure adopts an inward-facing conformation that is characterized by a cytosolic-facing cavity and solvent accessibility from the intracellular side to the substrate binding sites (Fig. 3A). Multiple polar residues on the surface of this cavity favor the recruitment of potassium and chloride ions (Fig. 3B). Above this cavity, several hydrophobic residues form a constriction that blocks extracellular solvent access (Fig. 3C). The extracellular gate is sealed by salt bridges and hydrogen bonds contributed by residues Arg^{140}, Glu^{222}, Asp^{575}, Lys^{485}, and Tyr^{486} (Fig. 3D). These interactions are expected to be disrupted when the extracellular gate opens, as observed in LeuT structures, where a salt bridge between Arg

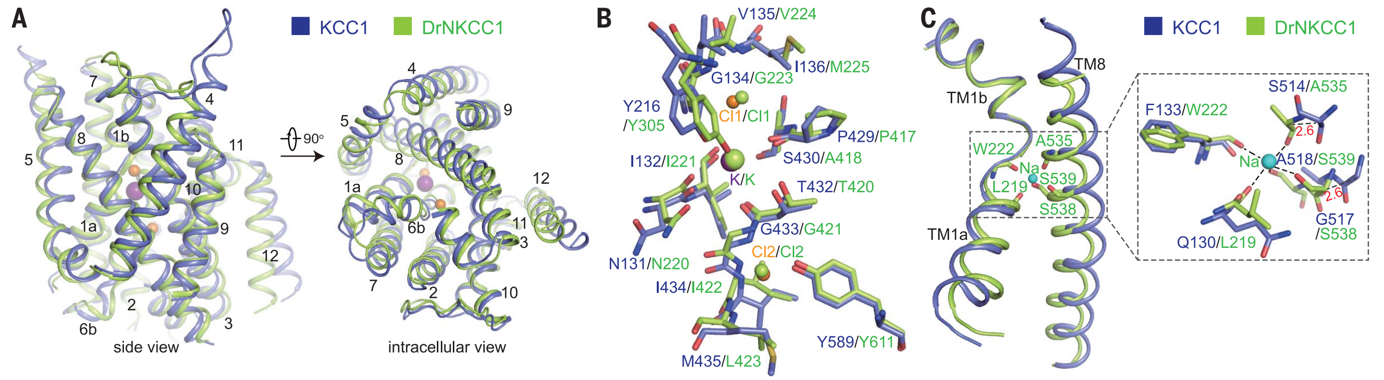


Fig. 4. Structural comparisons of human KCC1 and DrNKCC1. (**A**) The structural superimposition of KCC1 (blue) and DrNKCC1 (green) [Protein Data Bank (PDB) 6NPL] in the TMD. For clarity, large extracellular or intracellular loops are not shown. (**B**) The K^+ and Cl^- bind at similar positions in KCC1 (blue) and DrNKCC1 (green). In the KCC1 structure, K^+ and Cl^- are shown as purple and orange spheres, respectively, whereas in DrNKCC1, all three ions are shown as green spheres. (**C**) Compared with DrNKCC1 (green), KCC1 (blue) loses the Na^+ binding site. The Na^+ in the DrNKCC1 structure is shown as a cyan sphere. Red numbers show the distances in angstroms between Cα atoms of equivalent residues in KCC1 and DrNKCC1.

and Asp residues also participates in closing the extracellular gate (*27*).

On the surface of the TMD, two regions cover the extracellular gate. First, a long extracellular linker between TM7 and TM8 folds into two short helices (EL6 and EL7) and occludes the extracellular pathway by packing tightly against TM1b and TM7 on one side and TM3 and TM8 on the other side. Second, two antiparallel loops (pre-6a loop and post-EL5 loop) from the ECD form a lid-like structural unit on top of TM6a and TM1b (Fig. 3D). EL6, EL7, the pre-6a loop, and the post-EL5 loop are tightly sandwiched by the TMD and the ECD (Fig. 3E). The packing interactions imposed on TM1b and TM6a by EL6, EL7, the pre-6a loop, and the post-EL5 loop are further stabilized by dimerization of the ECD in a relatively rigid architecture. Therefore, the ECD probably modulates KCC1 activity by constraining conformational changes on TM1b and TM6a when the extracellular gate opens.

Structural comparisons of KCC1 and DrNKCC1 reveal the conservation and divergence among different CCC members. First, both KCC1 and DrNKCC1 adopt a LeuT-like fold in an inward-facing conformation, with a root-mean-square deviation of 1.5 Å over 389 Cα atoms between two structures in the TMD (Fig. 4A). Second, the identified S_K, S_{Cl1}, and S_{Cl2} sites in KCC1 align well with those in DrNKCC1 (Fig. 4B). In DrNKCC1, the Na^+ sits in the Na2 site that is conserved across many APC transporters (*16*, *17*); by contrast, in KCC1, the two Na^+ coordination Ser residues are replaced with Gly and Ala, accompanying a 2- to 3-Å shift of the TM8 backbone (Fig. 4C). With the loss of the Na^+ site in KCC1, the energy barrier for transporting the Cl^- at the S_{Cl2} site is unlikely to be crossed. Third, KCC1 and DrNKCC1 share a similar dimeric architecture with relatively loose packing, mediated by either lipid or detergent molecules in the TMD (fig. S18). The long ECD in KCC1 participates in the dimeric assembly, whereas the shorter CAP domain in DrNKCC1 does not.

The structural and functional characterization of KCC1 allows us to generate a K-Cl cotransport model (fig. S19). The KCC1 structure provides a framework for interpretation of disease-related mutations of other KCCs such as KCC2 and KCC3, whose mutations cause genetic disorders of the human nervous system (fig. S20) (*8*–*10*). The structural studies of human KCC1, together with zebrafish NKCC1, will promote drug development and disease treatment targeting human CCCs.

REFERENCES AND NOTES

1. S. C. Hebert, D. B. Mount, G. Gamba, *Pflugers Arch.* **447**, 580–593 (2004).
2. G. Gamba, *Physiol. Rev.* **85**, 423–493 (2005).
3. J. M. Russell, *Physiol. Rev.* **80**, 211–276 (2000).
4. K. B. Gagnon, E. Delpire, *Am. J. Physiol. Cell Physiol.* **304**, C693–C714 (2013).
5. C. Rivera *et al.*, *Nature* **397**, 251–255 (1999).
6. D. B. Simon *et al.*, *Nat. Genet.* **13**, 183–188 (1996).
7. D. B. Simon *et al.*, *Nat. Genet.* **12**, 24–30 (1996).
8. H. C. Howard *et al.*, *Nat. Genet.* **32**, 384–392 (2002).
9. K. T. Kahle *et al.*, *EMBO Rep.* **15**, 766–774 (2014).
10. T. Stödberg *et al.*, *Nat. Commun.* **6**, 8038 (2015).
11. S. N. Orlov, S. V. Koltsova, L. V. Kapilevich, N. O. Dulin, S. V. Gusakova, *Biochemistry* **79**, 1546–1561 (2014).
12. M. Gagnon *et al.*, *Nat. Med.* **19**, 1524–1528 (2013).
13. M. Puskarjov, K. T. Kahle, E. Ruusuvuori, K. Kaila, *Epilepsia* **55**, 806–818 (2014).
14. B. Chen *et al.*, *Cell* **174**, 521–535.e13 (2018).
15. S. Warmuth, I. Zimmermann, R. Dutzler, *Structure* **17**, 538–546 (2009).
16. T. A. Chew *et al.*, *Nature* **572**, 488–492 (2019).
17. A. Yamashita, S. K. Singh, T. Kawate, Y. Jin, E. Gouaux, *Nature* **437**, 215–223 (2005).
18. A. M. Hartmann *et al.*, *J. Biol. Chem.* **285**, 23994–24002 (2010).
19. S. Faham *et al.*, *Science* **321**, 810–814 (2008).
20. M. L. Moore-Hoon, R. J. Turner, *Biochemistry* **39**, 3718–3724 (2000).
21. S. Casula *et al.*, *J. Biol. Chem.* **276**, 41870–41878 (2001).
22. J. C. de Jong *et al.*, *J. Biol. Chem.* **278**, 24302–24307 (2003).
23. P. Blaesse *et al.*, *J. Neurosci.* **26**, 10407–10419 (2006).
24. S. Casula *et al.*, *Blood Cells Mol. Dis.* **42**, 233–240 (2009).
25. R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, R. MacKinnon, *Nature* **415**, 287–294 (2002).
26. M. L. Jennings, M. F. Adame, *Am. J. Physiol. Cell Physiol.* **281**, C825–C832 (2001).
27. H. Krishnamurthy, E. Gouaux, *Nature* **481**, 469–474 (2012).

ACKNOWLEDGMENTS

We thank Y. Jiang's support for the initiation of this project. We are grateful for X. Zhang's help with data processing, B. Liu's assistance with protein expression, and F. Yang's support with the Tl^+ uptake assay. Single-particle cryo-EM data were collected at the Center of Cryo-Electron Microscopy at Zhejiang University, UT Southwestern Medical Center Cryo-Electron Microscopy Facility, and Cryo-EM Platform of Peking University. We thank X. Zhang, N. Gao, and N. Li for support with facility access and data acquisition. **Funding:** This work was supported in part by the Ministry of Science and Technology (2018YFA0508100 to J.G. and 2016YFA0500404 to S.Y.), the National Natural Science Foundation of China (31870724 to J.G.; 31525001 and 31430019 to S.Y.; and 11722434 and 11874319 to J.L.), Zhejiang Provincial Natural Science Foundation (LR19C050002 to J.G. and LR16H0900001 to W.Y.), and the Fundamental Research Funds for the Central Universities (to J.G. and S.Y.). S.C. is supported by grants from laboratory and equipment management, Zhejiang University (SJS201814). X.B. is supported by Welch Foundation grant I-1944, the Cancer Prevention and Research Initiative of Texas and the Virginia Murchison Linthicum Scholar in Medical Research fund. E.D. is supported by NIH grants DK093501 and DK110375 and by Leducq Foundation grant 17CVD05. **Author contributions:** J.G., X.B., and S.Y. conceived and designed this project; S.L., J.G., C.Z., and F.W. prepared the samples; S.C., X.B., J.G., S.L., L.X., and S.Y. performed data acquisition, image processing, and structure determination; S.L., M.Z., and W.Y. performed the cell imaging assay; B.H. and J.L. performed modeling and simulation; and E.D. performed the K^+ influx assay. All authors participated in the data analysis and manuscript preparation. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Structure coordinates and cryo-EM density maps have been deposited in the Protein Data Bank under accession numbers 6KKR and EMD-0701 for KCC1 in 150 mM KCl and GDN; 6KKT and EMD-0702 for KCC1 in 150 mM KCl and nanodisc; and 6KKU and EMD-0703 for KCC1 in 150 mM NaCl and GDN.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/505/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S20
Table S1
References (*28*–*47*)

View/request a protocol for this paper from *Bio-protocol*.

13 June 2019; accepted 26 September 2019
10.1126/science.aay3129

SPIN PHYSICS

Coherent spin manipulation of individual atoms on a surface

Kai Yang[1]*, William Paul[1], Soo-Hyon Phark[1,2,3], Philip Willke[1,2,4], Yujeong Bae[1,2,4], Taeyoung Choi[2,3], Taner Esat[1,2,4], Arzhang Ardavan[5], Andreas J. Heinrich[2,3]*, Christopher P. Lutz[1]*

Achieving time-domain control of quantum states with atomic-scale spatial resolution in nanostructures is a long-term goal in quantum nanoscience and spintronics. Here, we demonstrate coherent spin rotations of individual atoms on a surface at the nanosecond time scale, using an all-electric scheme in a scanning tunneling microscope (STM). By modulating the atomically confined magnetic interaction between the STM tip and surface atoms, we drive quantum Rabi oscillations between spin-up and spin-down states in as little as ~20 nanoseconds. Ramsey fringes and spin echo signals allow us to understand and improve quantum coherence. We further demonstrate coherent operations on engineered atomic dimers. The coherent control of spins arranged with atomic precision provides a solid-state platform for quantum-state engineering and simulation of many-body systems.

Detecting and controlling the coherent dynamics of artificial spin systems with single-spin resolution could provide fundamental insight into quantum magnetism (*1*, *2*). A key technique for realizing coherent control is pulsed spin resonance, which traditionally uses short bursts of oscillating magnetic fields to induce transitions between selected quantum spin states (*3*). Coherent spin manipulation lies at the heart of spin-based quantum information processing and coherent spintronic devices (*4*). In the solid-state environment, coherent control of single spins has been performed on molecules in break junctions (*5*), quantum dots and dopants in semiconductors (*6*, *7*), and nitrogen-vacancy centers in diamond (*8*, *9*). It would be desirable to image and precisely modify the local environment and energy-level structures at the atomic scale.

Atomic-scale structures can be constructed and imaged with a scanning tunneling microscope (STM), which can probe spin-spin interactions (*10*, *11*), spin dynamics (*12*–*14*), and single-molecule motion (*15*, *16*). Notably, artificial nanostructures built by STM have developed into a fruitful testing ground for exploring quantum magnetism (*11*, *12*, *14*, *17*–*22*). Recently, continuous-wave (CW) electron spin resonance (ESR) was combined with spin-polarized STM to detect spin states with highly improved energy resolution (*22*–*24*).

The time evolution of a spin state can be accurately controlled by applying pulsed electromagnetic fields (*3*). In semiconductor heterostructures, coherent spin dynamics have been measured with STM by using femtosecond light pulses (*13*). Single-spin coherent control has not been achieved in STM because the spin manipulation time was long compared to the coherence time (*22*). Here, we demonstrate coherent spin manipulations at the single-atom scale, achieved by modulating the magnetic exchange interaction between the STM tip and the Ti atom on a surface using the time-varying electric field in the STM junction (Fig. 1A) (*25*).

The Ti atoms were separated from the Ag substrate with a bilayer MgO film and were most likely hydrogenated (supplementary materials and methods) (*24*). We refer to them simply as Ti. Each Ti atom has an electron spin $S = 1/2$ and was adsorbed either on top of a surface oxygen atom or at a bridge site between two oxygen atoms (*24*, *26*). The two spin states $|\downarrow\rangle$ and $|\uparrow\rangle$, with electron spin quantum number $m_s = \pm 1/2$, were separated by the Zeeman energy in the presence of an applied magnetic field (Fig. 1B). In addition to the direct-current (DC) voltage V_{DC}, a radiofrequency (RF) voltage was applied with an amplitude V_{RF} at the STM tunnel junction (*22*, *24*).

The magnetic field experienced by a Ti spin ($\mathbf{S}$) is the vector sum of the externally applied static field $\mathbf{B}_{ext}$ and the effective tip field, which, in the presence of V_{RF} with frequency f, has a static component $\mathbf{B}_{tip}$ and an oscillatory component $\Delta\mathbf{B}_{tip}\cos(2\pi f t)$. The Hamiltonian is given by (*25*)

$$H = \gamma\hbar[\mathbf{B} + \Delta\mathbf{B}_{tip}\cos(2\pi f t)]\cdot\mathbf{S} \qquad (1)$$

where γ is the gyromagnetic ratio and $\hbar$ is Planck's constant. $\mathbf{B} = \mathbf{B}_{ext} + \mathbf{B}_{tip}$ is the total static field. As long as $\Delta\mathbf{B}_{tip}$ is noncollinear with $\mathbf{B}$, it can drive transitions between $|\downarrow\rangle$ and $|\uparrow\rangle$ states when hf matches the Zeeman splitting set by $\mathbf{B}$ (Fig. 1B, bottom). This change in magnetic states is visible as a change in the steady-state current in the CW-ESR spectrum (Fig. 1C).

To access the fast coherent spin dynamics of a Ti atom, we applied a series of RF pulses at the resonant frequency to induce Rabi oscillations between the two spin states (Figs. 1A and 2A). Between pulses, the Ti spin was reinitialized to the $|\downarrow\rangle$ state by a spin-polarized current. The degree of polarization of the initial state was determined by the spin polarization of the tip (*27*). During the RF pulse, the Ti spin rotated coherently between the $|\downarrow\rangle$ and $|\uparrow\rangle$ states at the Rabi frequency Ω, so increasing the pulse width τ yielded an oscillatory current signal ΔI (Fig. 2, A and B). The exponential decay of ΔI gives the coherence time $T_2^{Rabi} \approx 40$ ns. When the RF frequency was detuned away from the Larmor frequency, Ω increased and the oscillation amplitude was reduced (fig. S10).

The Rabi oscillation of the Ti spin was read out magnetoresistively by using a homodyne detection technique (Fig. 2C and supplementary text section 4), a standard technique for coherent detection in physics and engineering (*28*, *29*). Here the asymmetric CW-ESR line shape is the result of homodyne demodulation (*26*). When the moment of the magnetic tip was canted with respect to the quantization axis of the Ti spin, the Larmor precession of the Ti spin yielded an RF conductance G_{RF}. The Rabi oscillations of the Ti spin modulated the amplitude of G_{RF} and were then extracted by measuring the time-average current due to V_{RF} (Fig. 2C). This homodyne detection yielded a better signal-to-noise ratio than measurements using DC pulses (supplementary text section 5), so homodyne detection results are reported below.

The spatial orientation of the tip magnetic moment was crucial for observing the Rabi oscillations (supplementary text section 6), because both the driving field and the amplitude of G_{RF} increased with the angle θ between the tip magnetic moment and $\mathbf{B}_{ext}$ (inset of Fig. 1C).

We gained additional control of Ω by adjusting the tip–atom distance d_{tip}, in addition to varying V_{RF}. Ω increased linearly with the setpoint current I_{DC} at a constant V_{DC} and V_{RF} (Fig. 2D). Because I_{DC} depended exponentially on d_{tip}, this result demonstrated the exponential dependence of Ω on d_{tip}, an important characteristic of exchange-field–driven ESR (*25*). We note that T_2^{Rabi} increased with larger d_{tip} (Fig. 2D), because one of the dominant decoherence sources was the tunneling current (supplementary text section 7) (*30*). For any given tip, T_2^{Rabi} and Ω scaled inversely with varying tip height. Thus, for a given tip and V_{DC}, the number of observable Rabi cycles was nearly independent of the tip height and could possibly be increased by using the magnetic dipolar coupling to the tip, or by placing a surface

[1]IBM Almaden Research Center, San Jose, CA 95120, USA. [2]Center for Quantum Nanoscience, Institute for Basic Science (IBS), Seoul 03760, Republic of Korea. [3]Department of Physics, Ewha Womans University, Seoul 03760, Republic of Korea. [4]Ewha Womans University, Seoul 03760, Republic of Korea. [5]CAESR, Clarendon Laboratory, Department of Physics, University of Oxford, Oxford OX1 3PU, UK.
*Corresponding author. Email: k.yang@ibm.com (K.Y.); heinrich.andreas@qns.science (A.J.H.); cplutz@us.ibm.com (C.P.L.)

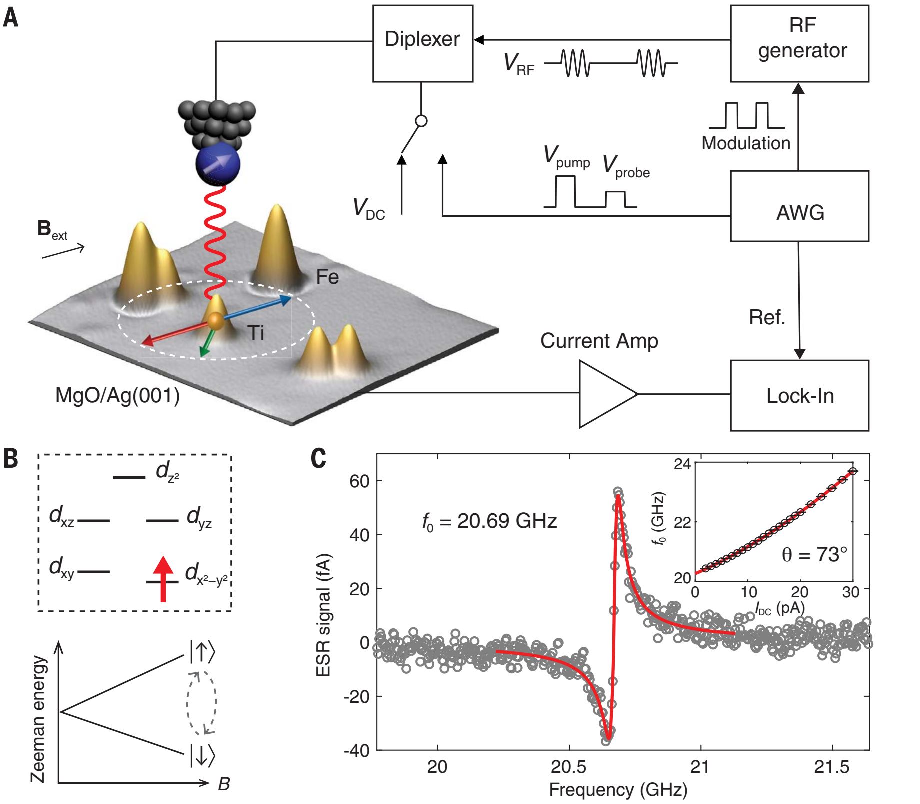


Fig. 1. An STM circuit to manipulate spins of individual atoms. (**A**) Schematic showing the low-temperature (T = 1.2 K) STM integrated with a radiofrequency (RF) generator and an arbitrary waveform generator (AWG), as well as an STM image (8.4 nm by 10 nm) of Ti and Fe atoms on a bilayer MgO on Ag(001) (setpoint: V_{DC} = 60 mV, I_{DC} = 10 pA). Exchange coupling (red wavy line) with the tip magnetic moment (white arrow) drives Ti spin resonance. The tip magnetoresistively senses the different Ti spin orientations, indicated by the red (|↓⟩ state), blue (|↑⟩ state), and green (superposition state) arrows. (**B**) Top: Orbital occupancy of the 3d^1 configuration. Bottom: Zeeman energy of the Ti spin. (**C**) CW-ESR spectrum of a single Ti atom, fitted to an asymmetric Lorentzian (eq. S14) (V_{DC} = 50 mV, I_{DC} = 5 pA, V_{RF} = 50 mV, B_{ext} = 0.82 T). Inset: resonant frequency f_0 as a function of I_{DC} at a constant V_{DC} = 50 mV. Fitting (red curve) to eq. S17 yields an angle θ = 73° between the tip's magnetic moment and $\mathbf{B}_{ext}$.

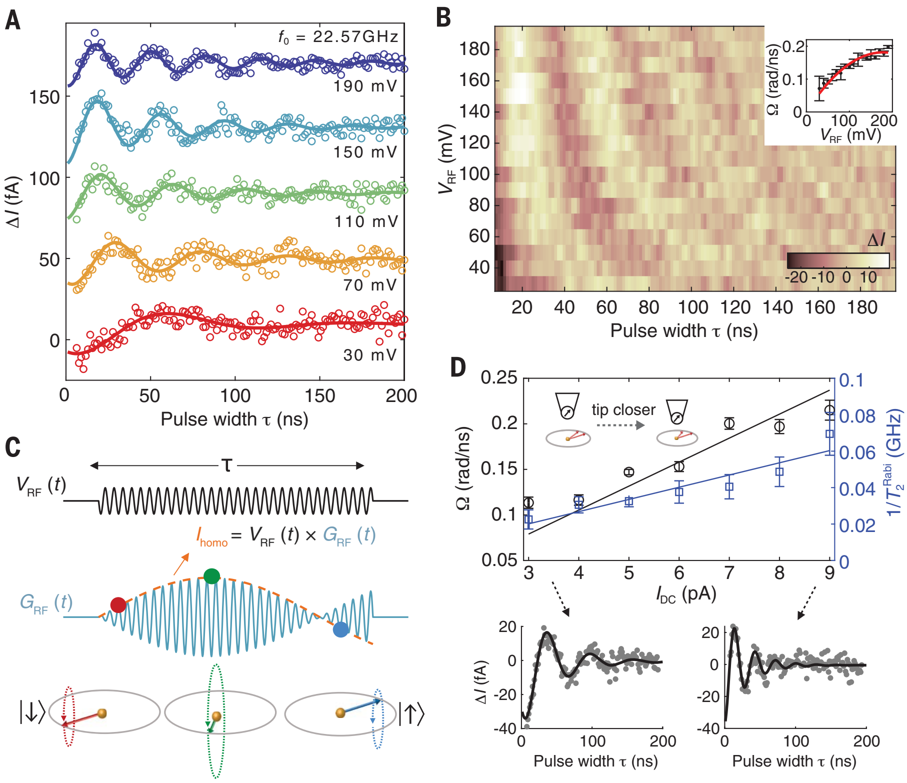
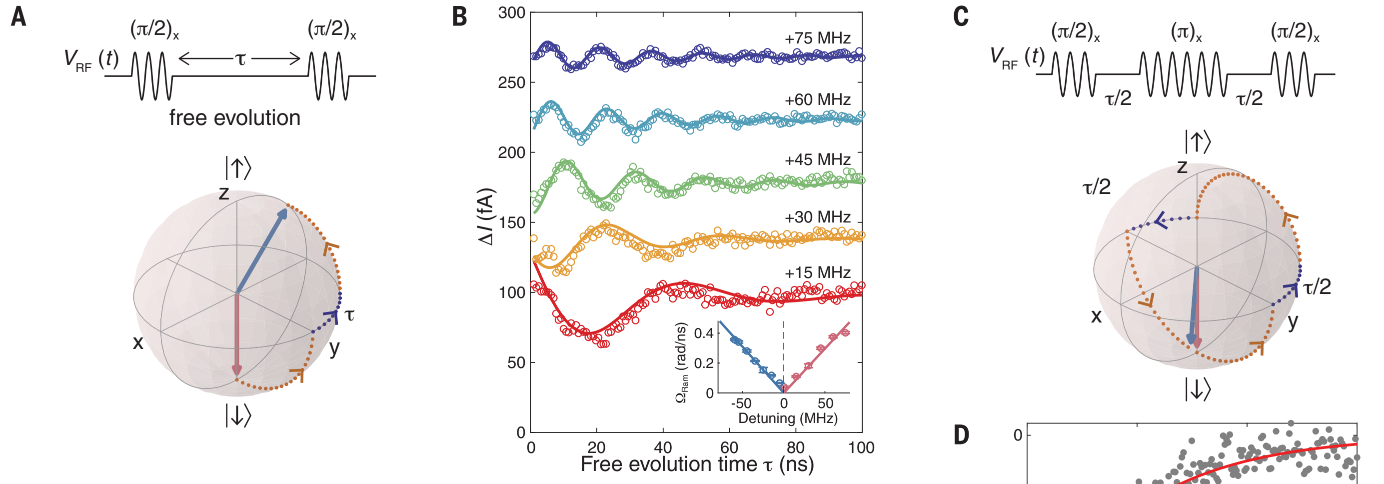


Fig. 2. Coherent control of a single Ti spin. (**A**) Rabi oscillations of a Ti spin at different values of V_{RF}, as indicated (V_{DC} = 60 mV, I_{DC} = 4 pA, B_{ext} = 0.90 T). The RF frequency is on resonance at f_0 = 22.57 GHz (fig. S6A). Solid curves are sinusoidal fits to the tunnel current signal ΔI. The data are offset vertically for better visibility. Each spectrum was measured for ~20 min. (**B**) Rabi oscillations represented in the color scale shown, over a fine grid of τ and V_{RF}. Inset: Ω at different values of V_{RF} fitted to a quadratic polynomial with zero offset. (**C**) Schematics showing the homodyne detection of the Rabi oscillations by measuring current I_{homo}, the time-averaged product of G_{RF} and V_{RF}. $G_{RF}(t)$ is the instantaneous tunnel conductance when $V_{RF}(t)$ is on. Pulse width τ was controlled by the AWG. (**D**) Ω and $1/T_2^{Rabi}$ at different tip–atom distances set by different I_{DC} at V_{DC} = 50 mV (V_{RF} = 80 mV, B_{ext} = 0.82 T). Solid lines are linear fits. Lower panels show the Rabi oscillations with fitting at I_{DC} = 3 pA (left) and 9 pA (right).



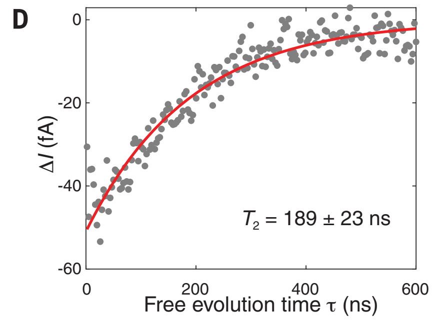


Fig. 3. Ramsey fringes and spin echo of a single Ti spin. (**A**) Pulse scheme (top) and Bloch sphere representation (bottom) of the spin evolution in the rotating frame for Ramsey fringes measurement. Red and blue arrows represent the Ti spin's initial or final state. Purple dots indicate free evolution between pulses. Orange dots represent a rotation about the *x* axis. (**B**) Ramsey signals and Ramsey fringe frequency Ω_{Ram} (inset) at different detuning frequencies $f - f_0$. (**C**) Pulse scheme (top) and Bloch sphere representation (bottom) for Hahn-echo measurement. (**D**) Spin-echo signals at different values of total free evolution time τ. The exponential fit (red curve) gives a coherence time of T_2 ~189 ns. Measured at V_{DC} = 50 mV, I_{DC} = 3 pA; V_{RF} = 190 mV, f_0 = 20.55 GHz, B_{ext} = 0.82 T.

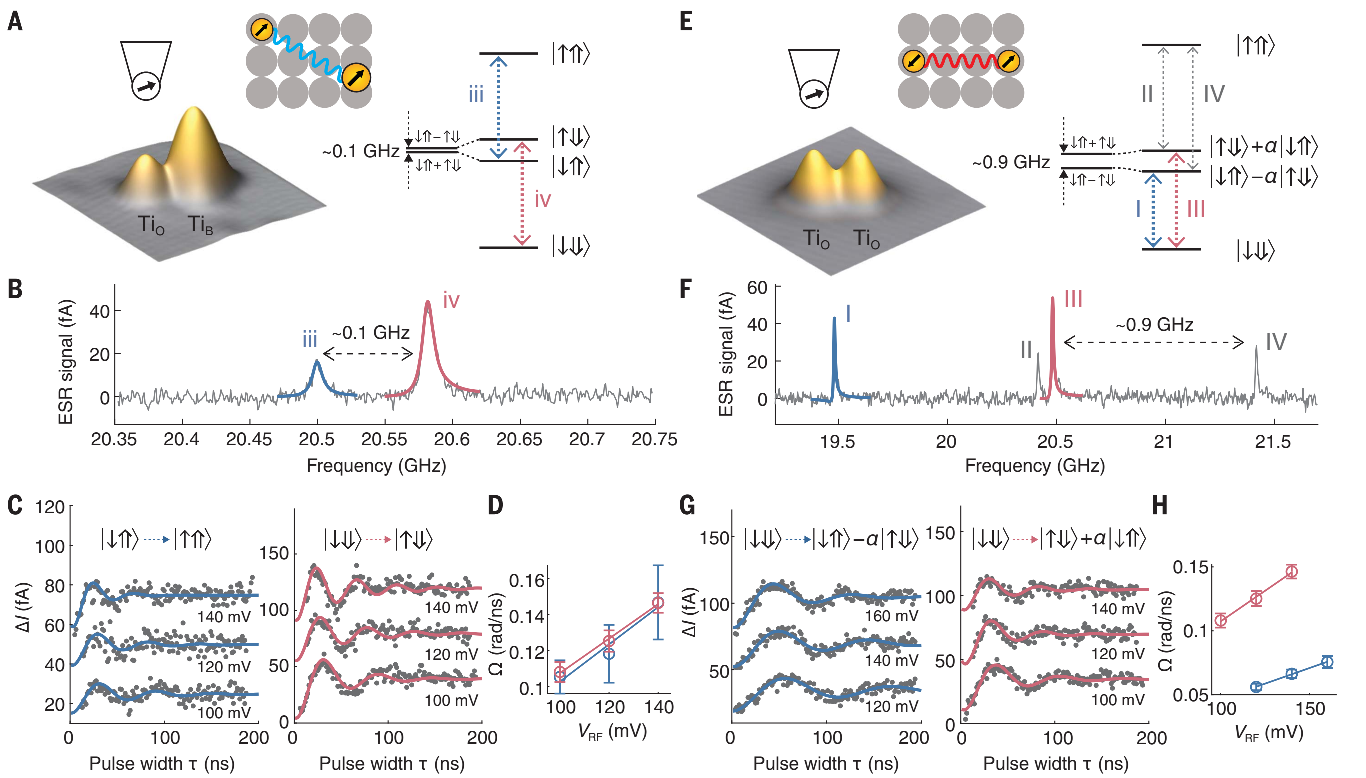


Fig. 4. Coherent control of spin dimers. (**A**) Left: Schematic showing an STM image (4 nm by 4 nm) of one oxygen-site and one bridge-site Ti atom (V_{DC} = 60 mV, I_{DC} = 30 pA), and the magnetic tip. Middle: Atom positions of Ti dimer on MgO. Gray circles represent oxygen atoms. Right: Energy level diagram with a ferromagnetic coupling of ~0.1 GHz. Dashed arrows indicate ESR transitions. (**B**) CW-ESR measured on the oxygen-site Ti atom (V_{DC} = 50 mV, I_{DC} = 3 pA, V_{RF} = 10 mV). (**C**) Pulsed ESR measured on the oxygen-site Ti atom at *f* = 20.50 GHz and 20.58 GHz, corresponding to the blue and red peaks in (B) (V_{DC} = 50 mV, I_{DC} = 3 pA), showing the coherent transitions from |↓⇑⟩ to |↑⇑⟩ (left), and from |↓⇓⟩ to |↑⇓⟩ (right). (**D**) Dependence of Ω on V_{RF} for transitions iii and iv. Ω is extracted from a sinusoidal fit in (C). (**E** to **H**) Same as in (A) to (D) except measured on two oxygen-site Ti atoms having a closer spacing, with an antiferromagnetic coupling of ~0.9 GHz. (G) The coherent transitions from |↓⇓⟩ to |↓⇑⟩−α|↑⇓⟩ (left) and from |↓⇓⟩ to |↑⇓⟩+α|↓⇑⟩ (right), where α ≈ 0.7. (H) Dependence of Ω on V_{RF} for transitions I and III. V_{RF} = 25 mV in (F). Measured at B_{ext} = 0.82 T.

spin nearby to supply the magnetic field gradient (*23*).

The coherence time of Rabi oscillations T_2^{Rabi} is, in principle, sensitive to (i) scattering electrons from tip or substrate (*31*), (ii) tunneling current arising through V_{DC} (*30*) and V_{RF}, (iii) variations of the resonant frequency caused by the slow-varying tip magnetic field, and (iv) variations in Ω caused by changes of tip–atom distance (Fig. 2D) or variations in V_{RF}. To understand and improve the spin coherence time, we applied multipulse spin manipulation sequences (Fig. 3).

Figure 3B shows the Ramsey fringes measured on a single Ti atom by applying two π/2 pulses separated by a time delay τ, where π represents an RF pulse that flips the spin from $|\downarrow\rangle$ to $|\uparrow\rangle$. The RF frequency was intentionally detuned from the resonance so that during the free evolution τ, a phase was accumulated between the two spin components. The time delay τ thus determined whether the spin returned to the $|\downarrow\rangle$ state or continued to evolve to the $|\uparrow\rangle$ state, leading to interference fringes. We directly controlled the Ramsey fringe frequency by choosing the frequency detuning $f - f_0$ (Fig. 3B, inset).

The decay of the Ramsey signal with τ also gave a coherence time T_2^*, which was about 40 ns. T_2^* was not substantially longer than T_2^{Rabi}, indicating that decoherence was dominated by sources other than current induced by V_{RF} or slow variations in Ω.

To improve the coherence time, we performed a Hahn-echo pulse sequence, consisting of a π pulse placed between two π/2 pulses: π/2–τ/2–π–τ/2–π/2, where τ is the duration of free precession (Fig. 3C). This sequence decoupled the Ti spin from any slowly changing magnetic fields (*32*). Here, the last π/2 pulse served as a homodyne detection pulse (*7*). The echo signal decayed exponentially with a time constant $T_2 \approx 189 \pm 23$ ns (Fig. 3D), which is several times longer than T_2^{Rabi} or T_2^*. This result indicates that the Ramsey coherence time T_2^* was limited by slow field variations that can be canceled by the echo technique. The main remaining decoherence source is the tunneling current and scattering electrons from the tip and substrate, and thus we would expect lower tunnel current and thicker MgO (*31*) to further extend T_2.

We further demonstrate coherent operations on coupled-spin states in designed atomic structures by assembling Ti atom pairs at chosen spacings (Fig. 4) (*24*). The spin Hamiltonian of a Ti spin dimer ($\mathbf{S}_1$ and $\mathbf{S}_2$) is

$$H = H_{\mathrm{Zee}} + H_{\mathrm{int}} \quad (2)$$

which consists of the Zeeman energy $H_{\mathrm{Zee}} = \gamma\hbar(B_{\mathrm{ext}} + B_{\mathrm{tip}})S_{1z} + \gamma\hbar B_{\mathrm{ext}}S_{2z}$ and the interaction energy $H_{\mathrm{int}} = J\mathbf{S}_1\cdot\mathbf{S}_2 + D(3S_{1z}S_{2z} - \mathbf{S}_1\cdot\mathbf{S}_2)$ (supplementary text section 11) (*24*). The Ti spin under the tip is $\mathbf{S}_1$. H_{int} contains an exchange coupling with strength *J* and dipolar coupling with strength *D* and produced a four-level energy spectrum with an unequal energy spacing, which could be characterized by CW-ESR (*24*, *26*).

The spin dimer in Fig. 4A had an interaction energy of ~0.1 GHz, much smaller than the Zeeman energy difference between the two spins induced by the tip field (~0.35 GHz). The energy eigenstates thus remained approximated as Zeeman product states ($|\downarrow\Downarrow\rangle$, $|\downarrow\Uparrow\rangle$, $|\uparrow\Downarrow\rangle$, and $|\uparrow\Uparrow\rangle$) (*24*). We performed controlled rotation of the spin under the tip (target spin), conditionally on the state of the other spin (control spin). When a pulse was applied at frequency f_{iv} (Fig. 4B), matching the energy spacing between the first and third energy levels (Fig. 4A, right panel), only the $|\downarrow\Downarrow\rangle \dashrightarrow |\uparrow\Downarrow\rangle$ transition was resonant with the RF field. Correspondingly, the RF field flipped the target spin if, and only if, the control spin was in state $|\Downarrow\rangle$ (Fig. 4C, right panel). Similarly, a π pulse at frequency f_{iii} flipped the target spin when the control spin was in state $|\Uparrow\rangle$. This capability provides a possible realization of a CNOT gate (*33*).

The spin dimer in Fig. 4E had a larger interaction of ~0.9 GHz, which overcame the magnetic asymmetry created by the tip and caused two of the quantum eigenstates to strongly mix to form states $|\downarrow\Uparrow\rangle - \alpha|\uparrow\Downarrow\rangle$ and $|\uparrow\Downarrow\rangle + \alpha|\downarrow\Uparrow\rangle$ with $\alpha \approx 0.7$ (*24*). The other two eigenstates remained the Zeeman product states (Fig. 4E, right panel). We coherently excited the system from the ground state $|\downarrow\Downarrow\rangle$ to either of the mixed states. When the RF pulse was tuned to transition I (Fig. 4F), $|\downarrow\Downarrow\rangle$ evolved into $|\downarrow\Uparrow\rangle - \alpha|\uparrow\Downarrow\rangle$ for a π pulse (Fig. 4G, left). In comparison, an RF pulse at the frequency of transition III coherently rotated $|\downarrow\Downarrow\rangle$ into $|\uparrow\Downarrow\rangle + \alpha|\downarrow\Uparrow\rangle$ (Fig. 4G, right). The Rabi frequency Ω grew in proportion to V_{RF} for these two transitions (Fig. 4H). For any given V_{RF}, the ratio of Ω for these two transitions depended only on the state mixing α (supplementary text section 11). As we increased the tip field, the state mixing was reduced (*24*), and thus the ratio of Ω between I and III became smaller (fig. S13). This tunable tip magnetic field thus provided control of the energy-level structures and Rabi rates.

Combining pulsed ESR with STM manipulation paves the way for the coherent excitation and detection of many-body states of artificial spin structures, such as topological quantum states of atomic chains (*34*). The exchange-field–driven coherent manipulation should be applicable to other solid-state spin systems (fig. S14) (*5*, *7*). The time-domain ESR spectroscopy could extend the sensing capabilities of a single-atom quantum sensor (*8*, *9*, *32*) and possibly also be used for quantum information applications using surface atoms (*33*, *35*).

REFERENCES AND NOTES

1. T. Fukuhara *et al.*, *Nature* **502**, 76–79 (2013).
2. C. Senko *et al.*, *Science* **345**, 430–433 (2014).
3. A. Abragam, B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions* (Oxford Univ. Press, Oxford, 2012).
4. S. Schlauderer *et al.*, *Nature* **569**, 383–387 (2019).
5. S. Thiele *et al.*, *Science* **344**, 1135–1138 (2014).
6. K. C. Nowack, F. H. L. Koppens, Y. V. Nazarov, L. M. K. Vandersypen, *Science* **318**, 1430–1433 (2007).
7. J. J. Pla *et al.*, *Nature* **489**, 541–545 (2012).
8. J. R. Maze *et al.*, *Nature* **455**, 644–647 (2008).
9. G. Balasubramanian *et al.*, *Nature* **455**, 648–651 (2008).
10. G. Czap *et al.*, *Science* **364**, 670–673 (2019).
11. C. F. Hirjibehedin, C. P. Lutz, A. J. Heinrich, *Science* **312**, 1021–1024 (2006).
12. S. Loth, M. Etzkorn, C. P. Lutz, D. M. Eigler, A. J. Heinrich, *Science* **329**, 1628–1630 (2010).
13. S. Yoshida *et al.*, *Nat. Nanotechnol.* **9**, 588–593 (2014).
14. A. A. Khajetoorians *et al.*, *Science* **339**, 55–59 (2013).
15. T. L. Cocker, D. Peller, P. Yu, J. Repp, R. Huber, *Nature* **539**, 263–267 (2016).
16. S. Li, S. Chen, J. Li, R. Wu, W. Ho, *Phys. Rev. Lett.* **119**, 176002 (2017).
17. A. A. Khajetoorians *et al.*, *Nat. Phys.* **8**, 497–503 (2012).
18. R. Toskovic *et al.*, *Nat. Phys.* **12**, 656–660 (2016).
19. R. Drost, T. Ojanen, A. Harju, P. Liljeroth, *Nat. Phys.* **13**, 668–671 (2017).
20. A. J. Heinrich, J. A. Gupta, C. P. Lutz, D. M. Eigler, *Science* **306**, 466–469 (2004).
21. F. Meier, L. Zhou, J. Wiebe, R. Wiesendanger, *Science* **320**, 82–86 (2008).
22. S. Baumann *et al.*, *Science* **350**, 417–420 (2015).
23. F. D. Natterer *et al.*, *Nature* **543**, 226–228 (2017).
24. K. Yang *et al.*, *Phys. Rev. Lett.* **119**, 227206 (2017).
25. K. Yang *et al.*, *Phys. Rev. Lett.* **122**, 227203 (2019).
26. Y. Bae *et al.*, *Sci. Adv.* **4**, eaau4159 (2018).
27. S. Loth *et al.*, *Nat. Phys.* **6**, 340–344 (2010).
28. C. Gross *et al.*, *Nature* **480**, 219–223 (2011).
29. N. Lee *et al.*, *Science* **332**, 330–333 (2011).
30. P. Willke *et al.*, *Sci. Adv.* **4**, q1543 (2018).
31. W. Paul *et al.*, *Nat. Phys.* **13**, 403–407 (2017).
32. C. L. Degen, F. Reinhard, P. Cappellaro, *Rev. Mod. Phys.* **89**, 035002 (2017).
33. A. Gaita-Ariño, F. Luis, S. Hill, E. Coronado, *Nat. Chem.* **11**, 301–309 (2019).
34. S. Nadj-Perge *et al.*, *Science* **346**, 602–607 (2014).
35. D. M. Zajac *et al.*, *Science* **359**, 439–442 (2018).

ACKNOWLEDGMENTS

We thank B. Melior for expert technical assistance. **Funding:** We acknowledge financial support from the Office of Naval Research. S.-H.P., P.W., Y.B., T.C., T.E., and A.J.H. acknowledge support from the Institute for Basic Science (IBS-R027-D1). A.A. acknowledges support from the Engineering and Physical Sciences Research Council (EP/L011972/1 and EP/P000479/1) and the QuantERA European Project SUMO. **Author contributions:** K.Y. and C.P.L. designed the experiment. K.Y., W.P., S-H.P., P.W., Y.B., T.C., and T.E. performed STM measurements. K.Y. and C.P.L. performed the analysis and wrote the manuscript with help from all authors. All authors discussed the results and edited the manuscript. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the supplementary materials. Additional data related to this paper may be requested from the authors.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/366/6464/509/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
References (*36*–*46*)

9 July 2019; accepted 1 October 2019
10.1126/science.aay6779

Steven Chu
AAAS President
William R. Kenan, Jr.
Professor of Physics
Stanford University

President's Address

Dear Colleagues:

Scientific and technological research have been invaluable in developing ways of improving life on Earth. The 2020 AAAS Annual Meeting theme, Envisioning Tomorrow's Earth, considers how efforts and advances in science and technology can respond to the new challenges faced by society.

On behalf of the AAAS Board of Directors, I urge you to join us in Seattle from February 13-16, 2020, where this theme will be considered from interdisciplinary scientific perspectives. The AAAS Annual Meeting is the premier event at which you can network with future collaborators from different disciplines. We look forward to seeing you in Seattle.

Join us in Seattle, Washington

Connect with scientific and technology professionals, policymakers, funders, educators, students, and members of the media. Together, explore the many ways that research and innovation are reacting to the challenges of the natural and built word:

- Participate in some of the 140+ scientific sessions in 11 disciplinary tracks covering advances in research and policy
- Attend workshops aimed at developing your career
- Take advantage of networking opportunities with new and familiar collaborators

Face-to-face meetings provide unique and unparalleled opportunities for engagement and networking that, as of yet, cannot be re-created virtually. AAAS recognizes the environmental impact of its and other scientific gatherings, and we are constantly examining ways to mitigate this impact. See the growing list of actions at **aaas.org/meetings/ecoefforts**, and send your own suggestions to **meetings@aaas.org**.

CONNECT WITH US!

@AAASmeetings
#AAASmtg

/AAAS.Science

aaas.org/meetings

Sponsors As of September 2019

AAAS, publisher of *Science*, thanks the sponsors and supporters of the 2020 Annual Meeting.

UK Research and Innovation

European Union

Johnson & Johnson INNOVATION

FRED HUTCH CURES START HERE®

Arizona State University

Topical Lectures

AI Advances and Aspirations

ERIC HORVITZ
Technical Fellow and Director, Microsoft Research

ASTRIAGraph: A Crowdsourced Knowledge Graph for Space Domain Awareness and Prediction

MORIBA JAH
Associate Professor of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin

Quantum Cloud Computing

ELHAM KASHEFI
Personal Chair in Quantum Computing, The University of Edinburgh

What Place for Solar Geoengineering and Carbon Removal in Climate Strategy?

DAVID KEITH
Gordon McKay Professor of Applied Physics, Harvard University

Baby Brains and our Neuro-Futures

PATRICIA KUHL
Professor and Co-Director, Institute for Learning & Brain Sciences, University of Washington

Seas the Day: Science Driving Ocean and Climate Solutions

THE HONORABLE JANE LUBCHENCO
Distinguished Professor, Oregon State University

The Second Antibiotic Era

MARYN MCKENNA
Journalist

Transgender Children

KRISTINA OLSON
Professor of Psychology, University of Washington

SARTON MEMORIAL LECTURE IN THE HISTORY AND PHILOSOPHY OF SCIENCE

The Indigenous/Science Project: Collaborative Practice as Witnessing

ALISON WYLIE
Professor, Department of Philosophy, University of British Columbia

Milestones Celebrations

19th Amendment

In August 1920, Congress ratified the 19th Amendment. The legal validation for women's voting rights has resulted, directly or indirectly, in the great advances women and other underrepresented groups have made in all spheres of public and private life.

Science—The Endless Frontier Report

In 1945, engineer Vannevar Bush created this extensive report for then-president Franklin D. Roosevelt. It provided many justifications for the value of the scientific endeavor and the ways it can support U.S. policy, and has had global and wide-ranging impact in government funding for the sciences in the years since its publication.

Earth Day

In 1970, Senator Gaylord Nelson of Wisconsin established the first Earth Day, an occasion to raise awareness about environmental concerns for individuals from all walks of life. In doing so, he began a national conversation about how best to protect America's natural resources in the years to come.

Updated program information will be posted on **aaas.org/meetings** as it becomes available.

Symposia

Advocating for the Future

ADAPTING TO CLIMATE CHANGE: IMPROVING CONNECTIONS BETWEEN SCIENCE AND SOCIETY
Organized by Amy Snover and Heidi Roop, University of Washington, Seattle, WA

ANIMAL CONSERVATION IN AFRICA: THE INTERSECTION OF SCIENCE AND SOCIETY
Organized by Bryan Ray and Sacha Vignieri, AAAS, Washington, DC

CULTURAL CONNECTIONS: COMMUNICATING SCIENCE TO COMMUNITIES AND CONGRESS
Organized by Bryn Nelson, Seattle, WA

DETECTING, COMBATING, AND IDENTIFYING DIS AND MIS-INFORMATION
Organized by Nadya Bliss, Arizona State University, Tempe, AZ

LEARNING THROUGH CITIZEN SCIENCE: ENHANCING OPPORTUNITIES BY DESIGN
Organized by Lekelia Jenkins, Arizona State University, Tempe, AZ

LOCAL SOLUTIONS TO GLOBAL CHALLENGES: A COMMUNITY-BASED APPROACH TO INNOVATION
Organized by Karel Haegeman, European Commission Joint Research Centre, Sevilla, Spain

PUBLIC TRUST IN SCIENCE: STRENGTH AND SKEPTICISM
Organized by Cary Funk, Pew Research Center, Washington, DC

SAVING SCIENCE JOURNALISM: ACTIONS FOR SCIENCE COMMUNICATION RESEARCHERS
Organized by Sue Ellen McCann, KQED, Inc., San Francisco, CA; Asheley Landrum, Texas Tech University, Lubbock, TX

SCIENCE ADVOCACY: LOBBYING FOR EVIDENCE
Organized by Sandra D. Mitchell, University of Pittsburgh, PA

SCIENCE COMMUNICATION STRATEGIES FOR BUILDING RELATIONSHIPS WITH THE PUBLIC
Organized by Elizabeth A. McCullough, Pacific Science Center, Seattle, WA

SCIENCE IS POLITICAL, NOT PARTISAN: BEST PRACTICES FOR SCIENCE POLICY ADVOCATES
Organized by Holly Mayton, National Science Policy Network, Virginia Beach, VA

THE REPRODUCIBILITY REVOLUTION: IMPACTS ON SCIENCE, JOURNALISM, AND SOCIETY
Organized by Simine Vazire, University of California, Davis, CA

WOMEN'S LEADERSHIP IN THE GLOBAL SCIENCE SYSTEM
Organized by Roseanne Diab, GenderInSITE, Trieste, Italy

Biomedical Futures

AVERTING CATASTROPHIC BIOLOGICAL INCIDENTS IN THE FUTURE
Organized by Kristin Omberg and Karen Taylor, Pacific Northwest National Laboratory, Richland, WA

BIOINFORMATICS AND AI: INNOVATIVE APPROACHES TO RESEARCH AND DATA STORAGE
Organized by Paul Whaley, Lancaster University, United Kingdom

THE BIOLOGICAL ENGINEERING REVOLUTION: STRATEGIES FOR A SUSTAINABLE SCALE-UP
Organized by Philip Shapira, Georgia Institute of Technology, Atlanta, GA; Carrie Cizauskas, Zymergen Inc., Emeryville, CA

BIOLOGICAL NORMALCY AND HUMAN VARIATION: THREE CASE STUDIES
Organized by Andrea Wiley, Indiana University, Bloomington, IN

DYNAMIC RNA MODIFICATIONS: ROLES IN DISEASE AND ENVIRONMENTAL RESPONSE
Organized by Frederick Tyson, National Institute of Environmental Health Sciences, Durham, NC

SYNTHETIC BIOLOGY AND THE NEXT-GENERATION BIOECONOMY
Organized by Emily Aurand, Engineering Biology Research Consortium, Emeryville, CA

EQUITY IN GENOMIC MEDICINE: GIVING VOICE TO INDIGENOUS GENETIC VARIATION
Organized by Wyeth Wasserman, University of British Columbia, Vancouver, Canada

GREEN TOXICOLOGY
Organized by Alexandra Maertens and Thomas Hartung, Johns Hopkins University, Baltimore, MD

HUMAN CELL ATLAS: TRANSFORMING BIOLOGY AND HEALTHCARE
Organized by Emily Mobley, Wellcome Sanger Institute, Cambridge, United Kingdom

LIVING DONORS: HOW VOLUNTEERS GIVE TISSUES AND TIME TO ADVANCE SCIENCE
Organized by Kathy Richmond, The Paul G. Allen Frontiers Group, Seattle, WA; Jennifer Pawlosky, Allen Institute, Seattle, WA

MOLECULAR PROBES AND IMAGING: EXPLORING THE INNER UNIVERSE
Organized by Jonathan Bagger and Paul Schaffer, TRIUMF, Vancouver, Canada

ORAL CANCER: USING THE HPV VACCINE FOR PREVENTION
Organized by Jacques E. Nor, University of Michigan, Ann Arbor, MI

PARTICLE THERAPY: PERSPECTIVES FROM PHYSICS, MEDICINE, AND ECONOMICS
Organized by Charles Clark, Joint Quantum Institute, Gaithersburg, MD; Anice Anderson, Private Engineering Consulting, Carmel, IN

PHARMACEUTICAL MANUFACTURING: SUSTAINABLE INNOVATIONS TO PRODUCE ZERO WASTE
Organized by Donna Huryn, University of Pittsburgh, PA; William Beck, University of Illinois, Chicago, IL

SEQUENCING THE GENOMES OF LIFE: PROTECTING BIODIVERSITY AND SUSTAINING SOCIETY
Organized by Emily Mobley, Wellcome Sanger Institute, Cambridge, United Kingdom

SYNTHETIC BIOLOGY: DIGITAL DESIGN OF LIVING SYSTEMS
Organized by Juerg Brunnschweiler and Marianne Lucien, ETH Zürich, Switzerland

Climate Futures

AI FOR CLIMATE CHANGE ADAPTATION
Organized by So-Min Cheong, University of Kansas, Lawrence, KS

BOUNDARY SPANNING: ENVIRONMENTAL SCIENCE AND DECISION-MAKING
Organized by Kei Koizumi, Shanghai, China

CARBON ECONOMY: CONSIDERATIONS FOR THE UNITED STATES AND PACIFIC RIM
Organized by Naoki Saito, RIKEN, Yokohama, Japan; Steven Collins, University of Washington, Bothell, WA

CHINA'S ROAD TOWARDS ECOLOGICAL CIVILIZATION AND ENVIRONMENTAL SUSTAINABILITY
Organized by Fahu Chen, Chinese Academy of Sciences, Beijing, China

CLIMATE AND MARINE ECOSYSTEMS: EXPLORATION, PREDICTION, AND PROTECTION
Organized by Heather Benway, Woods Hole Oceanographic Institution, MA

CLIMATE CHANGE AND CLIMATE EXTREMES: PRECIPITATION AND PROJECTIONS
Organized by Tianjun Zhou, Chinese Academy of Sciences, Beijing, China

CLIMATE CHANGE MITIGATION: THE ROLES OF INDUSTRY AND THE PRIVATE SECTOR
Organized by Murray Hitzman, University College Dublin, Belfield, Ireland

CLIMATE CHANGE SOLUTIONS: CAPTURING GREENHOUSE GAS EMISSIONS IN LAND
Organized by Sara Nichols, University of California, Davis, CA; Lifang Chiang, University of California Office of the President, Oakland, CA

FOSSIL FUEL PRODUCTION AND CLIMATE CHANGE: ALIGNING GOALS AND POLICIES
Organized by Georgia Piggot, Stockholm Environment Institute, Seattle, WA; Gretchen Goldman, Union of Concerned Scientists, Washington, DC

THE FUTURE OF EARTH'S CLIMATE: A WORLD OF EXTREMES
Organized by Anne Crasner, European Commission Executive Agency for Small and Medium-sized Enterprises, Brussels, Belgium

THE FUTURE OF SUSTAINABILITY SCIENCE IN THE UNITED STATES
Organized by Josh Tewksbury, University of Colorado and Colorado State University, Boulder, CO; Judit Ungvari-Martin, National Science Foundation, Alexandria, VA

THE GLOBAL CARBON CYCLE: IMPLICATIONS FOR LIFE BEYOND THE 21ST CENTURY
Organized by Forrest Hoffman, Oak Ridge National Laboratory, Oak Ridge, TN; Abigail Swann, University of Washington, Seattle, WA

IMPLEMENTING SUSTAINABLE SOLUTIONS IN A CHANGING WORLD
Organized by Julia Fahrenkamp, Science International, Cambridge, United Kingdom

IMPROVED WILDFIRE RESILIENCE: LEVERAGING SCIENCE AND TECHNOLOGY
Organized by Sarah Brady and Teresa Feo, California Council on Science and Technology, Sacramento, CA

OIL SPILLS: ADVANCES IN MITIGATING IMPACT AND CONSEQUENCES
Organized by Ken Halanych, Auburn University, Auburn, AL; Rita R. Colwell, University of Maryland, College Park, MD

SCIENCE DURING CRISIS: BEST PRACTICES, RESEARCH NEEDS, AND POLICY PRIORITIES
Organized by John Randell and Amanda Vernon, American Academy of Arts and Sciences, Cambridge, MA

USING MARINE ENERGY TO CONFRONT GLOBAL CHALLENGES
Organized by Gail Mattson, Brookhaven National Laboratory, Upton, NY

Digital Futures

COMMUNICATION THROUGH TOUCH: FROM BABIES TO SOCIAL ROBOTS
Organized by Andrew Meltzoff, University of Washington, Seattle, WA

THE DIGITAL TRANSFORMATION OF SOCIETIES: ADDRESSING TECHNOLOGICAL DISRUPTION
Organized by Jeffrey Alexander, RTI International, Rockville, MD; Eswaran Subrahmanian, Carnegie Mellon University, Pittsburgh, PA

GERRYMANDERING AND MATHEMATICS: REDISTRICTING THE NATION
Organized by Karen Saxe, American Mathematical Society, Washington, DC

HOW TO COUNTER INTERSECTIONAL BIASES IN SOCIAL MEDIA
Organized by Rochelle Diamond, National Organization of Gay and Lesbian Scientists and Technical Professionals, Pasadena, CA; Jon Pincus, A Change Is Coming, Bellevue, WA

READING EMOTIONS FROM FACIAL EXPRESSIONS: IMPLICATIONS FOR TECHNOLOGY AND HEALTH
Organized by Ralph Adolphs, California Institute of Technology, Pasadena, CA

RESILIENCE IN THE DIGITAL AGE
Organized by Christiane Rousseau, University of Montreal, Canada; Fred Roberts, Rutgers University, Piscataway, NJ

ROBOTIC COMPANIONS AND THE FUTURE OF AI
Organized by Igor Linkov and Benjamin Trump, United States Army Engineer Research and Development Center, Concord, MA

SOCIAL MEDIA AND THE TRANSFORMATION OF SCIENCE ADVICE
Organized by Sarah Foxen and Chris Tyler, University College London, United Kingdom

Engineering the Future

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING: DESIGNING FOR SAFETY AND SECURITY
Organized by Bill Pike and Courtney Corley, Pacific Northwest National Laboratory, Richland, WA

ARTIFICIAL INTELLIGENCE RESEARCH: A COMMUNITY ROADMAP
Organized by Ann Schwartz Drobnis, Computing Community Consortium, Washington, DC

CATALYSTS FOR ENERGY STORAGE: INSPIRED BY NATURE, BUILT BY SCIENTISTS
Organized by Wendy Shaw and Aaron Appel, Pacific Northwest National Laboratory, Richland, WA

CLEAN AVIATION IN TOMORROW'S WORLD
Organized by Michael Kyriakopoulos and Andrea Gentili, European Commission, Brussels, Belgium

CONSTRUCTING AND PERCEIVING BEAUTY
Organized by Daphne Maurer, McMaster University, Hamilton, Canada

DETECTING LIFE AND EXTRATERRESTRIAL TECHNOLOGIES
Organized by Anthony J. Beasley, National Radio Astronomy Observatory, Charlottesville, VA

ENGINEERING BIOLOGY AND THE PROMOTION OF COMMUNITY SECURITY
Organized by Jeffrey Fortman, Engineering Biology Research Consortium, Emeryville, CA

EXOPLANETS: THE DEVELOPMENT OF CURRENT KNOWLEDGE
Organized by Heidi Hammel, Association of Universities for Research in Astronomy, Washington, DC

NEW APPROACHES TO FAIRNESS IN AUTOMATED DECISION MAKING
Organized by Sampath Kannan, University of Pennsylvania, Philadelphia, PA; Ann Schwartz Drobnis, Computing Community Consortium, Washington, DC

NEXT GENERATION COMPUTER HARDWARE
Organized by Ann Schwartz Drobnis, Computing Community Consortium, Washington, DC

SOLAR GEOENGINEERING RESEARCH AROUND THE GLOBE
Organized by David Keith and Mariia Belaia, Harvard University, Cambridge, MA

VIEWING TOMORROW'S EARTH FROM SPACE AND SURFACE
Organized by Jack Kaye, National Aeronautics and Space Administration, Washington, DC

Future Earth Systems

ARCTIC INFLUENCES ON SEVERE WINTER WEATHER
Organized by James Overland, National Oceanic and Atmospheric Administration, Seattle, WA

CONVERGENT SCIENCE: COMBATING AFRICA'S WILDLIFE CRISIS
Organized by Mary Ann Ottinger, University of Houston, TX; Meredith Gore, Michigan State University, East Lansing, MI

ECOSYSTEMS' TRUE VALUE: EUROPEAN EFFORTS TO MAP AND ACCOUNT FOR NATURE
Organized by Josefina Enfedaque, European Commission Research and Innovation Directorate-General, Brussels, Belgium; Joachim Maes, European Commission Joint Research Centre, Geel, Belgium

ENVISIONING OCEAN CLIMATE SOLUTIONS FOR THE NEXT GENERATION
Organized by Stephen Posner, University of Vermont, Burlington, VT; Heather Mannix, COMPASS, Silver Spring, MD

THE FUTURE OF WATER AND HUMAN DECISION-MAKING
Organized by Ruby Leung, Pacific Northwest National Laboratory, Richland, WA; Gary Geernaert, Department of Energy, Germantown, MD

GEOSCIENCE LITERACY AND COMMUNITY RESILIENCE
Organized by Cathryn A. Manduca, Carleton College, Northfield, MN

THE GLOBAL WATER CYCLE: UNDERSTANDING TODAY AND TOMORROW
Organized by Michael Dettinger, United States Geological Survey, Carson City, NV

GRASSLANDS AND SAVANNAS: HUMAN IMPERATIVES AND BIODIVERSITY CONSERVATION
Organized by Michael Hill, University of North Dakota, Farrer, Australia

IS THE COAST TOAST? CASCADIA MEGA-EARTHQUAKES, TSUNAMIS, AND POTENTIAL IMPACTS
Organized by Harold Tobin and Alison Duvall, University of Washington, Seattle, WA

MANAGING WATER: NEW TOOLS FOR SUSTAINABLE DEVELOPMENT
Organized by Sera Young, Northwestern University, Evanston, IL

MARINE MAMMAL HEATH ASSESSMENTS: INFORMING SCIENTISTS, POLICY, AND THE PUBLIC
Organized by Stephen A. Raverty, British Columbia Ministry of Agriculture, Abbotsford, Canada; Mike E. Grigg, National Institutes of Health, Bethesda, MD

NATURE CONSERVATION AND COMPUTATIONAL TECHNOLOGIES: EXPANDING THE SCALE
Organized by Daniel Rubenstein, Princeton University, NJ; Tanya Berger-Wolf, University of Illinois, Chicago, IL

NATURE REMADE: ENGINEERING LIFE FROM THE PAST TO FUTURE WORLDS
Organized by Christian Young, Alverno College, Milwaukee, WI; Michael Dietrich, University of Pittsburgh, PA

OCEAN OUTBREAKS ON A CHANGING PLANET
Organized by C. Drew Harvell, Cornell University, Ithaca, NY

THE REEF CRISIS IN EARTH'S FUTURE
Organized by Jere H. Lipps, University of California, Berkeley, CA

USING GEOSPATIAL DATA TO LOWER EMISSIONS
Organized by Conrad M. Albrecht and Sharathchandra Pankanti, IBM Research, Yorktown Heights, NY

WILDFIRE SMOKE AND PUBLIC HEALTH: THE SCIENCE AT THE NEXUS
Organized by Ian Gilmour, United States Environmental Protection Agency, Research Triangle Park, NC

Future Health Strategies

AN AGING WORLD: PITFALLS AND PROMISE
Organized by Phyllis Moen, University of Minnesota, Minneapolis, MN

BEETHOVEN AT 250 AND THE SCIENCE OF MUSIC: EMOTION, MEMORY, AND HEALTH
Organized by Daniel Levitin, Minerva Schools, San Francisco, CA

DIET FOR A SICK PLANET: REDUCING EMISSIONS AND THE BURDEN OF CHRONIC DISEASE
Organized by Lydia Zepeda, University of Wisconsin, Madison, Tacoma, WA; Sean B. Cash, Tufts University, Boston, MA

THE DRUG ABUSE CRISIS: KEY FINDINGS FROM THREE LANDMARK STUDIES
Organized by Linda Teplin, Northwestern University, Chicago, IL

HEALTH DISPARITIES: THE INTERSECTION OF PSYCHOLOGY, HISTORY, LAW, AND MEDICINE
Organized by Sophie Trawalter and Dayna Matthew, University of Virginia, Charlottesville, VA

HUMAN EMBRYO RESEARCH REVISITED: SCIENTIFIC, ETHICAL, LEGAL, AND SOCIAL ISSUES
Organized by Kenneth Evans, Rice University, Houston, TX

INFECTIOUS DISEASE FORECASTING: MODELS AND MACHINE LEARNING
Organized by John Drake, University of Georgia, Athens, GA

THE NEUROSCIENCE OF ADDICTION: POLICY CONSIDERATIONS
Organized by Marina Picciotto, Yale University, New Haven, CT

SUPPORTING THE WHOLE STUDENT: MENTAL HEALTH, SUBSTANCE ABUSE, AND WELL-BEING
Organized by Layne Scherer, The National Academies of Sciences, Engineering, and Medicine, Washington, DC

TACKLING SOCIAL RISK FACTORS, HEALTH, AND CARE
Organized by Arlene S. Ash, University of Massachusetts, Worcester, MA

TRACING HOW SCIENTIFIC INFRASTRUCTURE ACCELERATES DISCOVERY
Organized by Elizabeth Lyons, National Science Foundation, Alexandria, VA

TRANSFORMING GLOBAL AND PUBLIC HEALTH THROUGH NATURE
Organized by Usha Varanasi, National Oceanic and Atmospheric Administration, Seattle, WA; Joshua Lawler, University of Washington, Seattle, WA

Future Societal Ethics

50TH ANNIVERSARY OF THE VIETNAM DRAFT: LESSONS FOR THE SCIENCES
Organized by Tim Johnson, Willamette University, Salem, OR

ETHICAL CONCERNS WITH ADVANCES IN TECHNOLOGY AND GENETICS
Organized by Subrata Saha and Pamela Saha, University of Washington, Seattle, WA

THE CHANGING IDENTITY LANDSCAPE: MULTIRACIAL, INTERSEX, & TRANSGENDER PEOPLE
Organized by Susan Gelman, University of Michigan, Ann Arbor, MI; Kristina Olson, University of Washington, Seattle, WA

DEATH IN THE 21ST CENTURY: WHAT IS LEFT BEHIND
Organized by Robert O'Malley, AAAS, Washington, DC

ETHICAL ISSUES IN ARTIFICIAL INTELLIGENCE
Organized by Joseph Halpern, Cornell University, Ithaca, NY

ETHICAL RISKS OF VOICE TECHNOLOGY: A SOCIOLINGUISTIC PERSPECTIVE
Organized by Emily M. Bender, University of Washington, Seattle, WA

ETHNOGRAPHY, OBSERVATION, AND NATURAL HISTORY: TOOLS FOR ETHICAL SCIENCE
Organized by Michelle Bezanson, Santa Clara University, Santa Clara, CA

EVOLUTION OF CRIMINAL JUSTICE: HOW SCIENCE AND EVIDENCE-BASED POLICY INTERACT
Organized by Trisha Chakraborty, Department of Justice, Washington, DC

EVOLVING CONCEPTS OF SCIENTIFIC INTEGRITY AND PRACTICE
Organized by Jonathan Coopersmith, Texas A&M University, College Station, TX

HOW CONGRESS USES SCIENCE AND TECHNOLOGY FOR POLICY: EMERGING RESEARCH
Organized by Karen Akerlof, George Mason University, Fairfax, VA; Chris Tyler, University College London, United Kingdom

IMMIGRATION, CRIME, AND JUSTICE: A DATA-DRIVEN INVESTIGATION
Organized by William Pridemore, State University of New York, Albany, NY

IMPERILED CULTURAL HERITAGE: MITIGATING THREATS, CO-PRODUCING KNOWLEDGE
Organized by Alyne Delaney, Aalborg University, Denmark

IMPLICIT BIAS, EXPLICIT SCIENCE
Organized by Erin Heath, AAAS, Washington, DC

LEARNING FROM PROTACTILE DEAFBLIND COMMUNITIES: TOWARD A MORE TACTILE FUTURE
Organized by Terra Edwards, Saint Louis University, MO; Diane Brentari, University of Chicago, IL

POLITICAL ANIMALS: BEHAVIOR, KNOWLEDGE, REASON, AND TOMORROW'S POLICYMAKING
Organized by David Mair and Marton Hajdu, European Commission Joint Research Center, Brussels, Belgium

PSYCHOLOGICAL SCIENCE: LESSONS FOR THE LAW
Organized by Nora Newcombe, Temple University, Philadelphia, PA

RESPONDING TO CLIMATE CHANGE: SCIENCE, RELIGION, AND CULTURAL PRACTICES
Organized by Curtis L. Baxter and Lilah Sloane, AAAS, Washington, DC

USING DATA AND AI TO DISRUPT SEX TRAFFICKING
Organized by Barbara Mack, National Council of Juvenile and Family Court Judges, Seattle, WA

The Future of Food

CROP DIVERSIFICATION: ENSURING A SUSTAINABLE GLOBAL FOOD SUPPLY
Organized by Raul Zornoza, Technical University of Cartagena, Spain; Antoine Messean, National Institute for Agricultural Research, Thiverval-Grignon, France

FOOD OF THE FUTURE: DEVELOPING NEW, SUSTAINABLE, AND HEALTHY SOURCES
Organized by Jens Wilkinson, RIKEN, Saitama, Japan

FOOD SAFETY REGULATIONS: CARCINOGENS AND THE RELEVANCE OF THE DELANEY CLAUSE
Organized by Mansi Krishan, Danone North America, Louisville, CO; Lisa Navarro, Givaudan Flavors Corporation, Cincinnati, OH

HERBICIDE RESISTANCE: CATALYZING TRANSDISCIPLINARY SCIENCE
Organized by David Ervin, Portland State University, OR

IMPROVING FOOD SECURITY POLICIES: INNOVATIVE AGRICULTURE AND MARKET MONITORING
Organized by Felix Rembold, European Commission Joint Research Centre, Ispra, Italy; Inbal Becker Reshef, University of Maryland, College Park, MD

NOURISHING PEOPLE AND PLANET: BUILDING A BETTER, BALANCED FOOD SYSTEM FOR 2050
Organized by Catherine Woteki, Iowa State University, Ames, IA

SUSTAINABLE AGRICULTURE: LAUNCHING STANDARDIZED MICROBIAL ECOSYSTEMS
Organized by Trent Northen, Lawrence Berkeley National Laboratory, Berkeley, CA; Jo Handelsman, University of Wisconsin-Madison, WI

TOMORROW'S TABLE: PLANT GENETICS AND THE FUTURE OF FOOD
Organized by Pam Ronald, University of California, Davis, CA

USING COMPUTING TO SUSTAINABLY FEED A GROWING POPULATION
Organized by Shashi Shekhar, University of Minnesota, Minneapolis, MN; James Hodson, AI for Good Foundation, El Cerrito, CA

Transforming Future Learning

CITIZEN SCIENCE AND BIG DATA: FROM ENGAGEMENT TO ACTION
Organized by Julia K. Parrish, University of Washington, Seattle, WA

DIVERSITY, EQUITY, AND INCLUSION: RECOMMENDATIONS FROM PHYSICS AND ASTRONOMY
Organized by Arlene Modeste-Knowles and Philip W. Hammer, American Institute of Physics, College Park, MD

FINDING THE LOST EINSTEINS
Organized by Michael Feder, AAAS, Washington, DC

INCLUSION IN THE ACADEMY: EXPLORING IDENTITY AND EQUITY IN THE STEM COMMUNITY
Organized by Lina Dahlberg and Robin Kodner, Western Washington University, Bellingham, WA

INCLUSIVITY AND EQUITY IN COURSE-BASED UNDERGRADUATE RESEARCH EXPERIENCES
Organized by Jeffrey Olimpo, University of Texas at El Paso, TX

MENTORSHIP: TOWARD A CULTURE OF EFFECTIVENESS AND INCLUSIVITY
Organized by Maria Lund Dahlberg, National Academy of Sciences, Engineering, and Medicine, Washington, DC

PUBLIC ENGAGEMENT AT UNIVERSITIES: PATHWAYS FOR INSTITUTIONAL SUPPORT
Organized by John Meyer, University of Washington, Seattle, WA; Emily Cloyd, AAAS, Washington, DC

STEM RESEARCH EXPERIENCES FOR HIGH SCHOOL STUDENTS
Organized by Raja GuhaThakurta, University of California, Santa Cruz, CA; Or Graur, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA

STRENGTHENING SUSTAINABILITY PROGRAMS AND CURRICULA IN HIGHER EDUCATION
Organized by Lida Beninson, National Academies of Sciences, Engineering, and Medicine, Washington, DC

TRANSFORMING HIGHER EDUCATION CULTURE: COORDINATING REFORM WITH AGENTS OF CHANGE
Organized by Barbara Natalizio, The Rita Allen Foundation, Princeton, NJ; Erica Kimmerling, American Academy of Arts and Sciences, Cambridge, MA

WOMEN IN STEMM: ADDRESSING UNDERREPRESENTATION
Organized by Alex Helman and Ashley Bear, National Academies of Sciences, Engineering, and Medicine, Washington, DC

Urban Futures

BIOLOGICAL INVASION FORECASTING: UNITING CLIMATE CHANGE, TRANSPORT, AND TRADE
Organized by Erin Grey, Governors State University, University Park, IL; David Lodge, Cornell University, Ithaca, NY

FACTORIES REIMAGINED: MAKING INDUSTRIAL WORK MORE APPEALING
Organized by Erastos Filos, European Commission Research and Innovation Directorate, Brussels, Belgium; Eija Kaasinen, VTT Technical Research Centre of Finland Ltd., Tampere, Finland

NATURAL HISTORY COLLECTIONS AND BIODIVERSITY'S FUTURE
Organized by Keegan Sawyer and Audrey Thevenon, National Academies of Sciences, Engineering, and Medicine, Washington DC

SMART NEW ENERGY VEHICLES AND INTELLIGENT TRANSPORTATION
Organized by Yingjun Qiao, Chinese Academy of Engineering, Beijing, China

SOCIALLY INTEGRATIVE CITIES: PAVING THE WAY TO URBAN SUSTAINABILITY
Organized by Bernhard Müller, Leibniz Institute of Ecological Urban and Regional Development, Dresden, Germany

URBAN PAVING IS GOING PLACES
Organized by Cesare Sangiorgi, University of Bologna, Italy; Ioannis Bitsios, European Commission Research Executive Agency, Brussels, Belgium

URBAN RESILIENCE AND EMERGENCY RESPONSE: THE CLIMATE CHANGE PERSPECTIVE
Organized by Julie Dirwimmer, Fonds de recherche du Québec, Montréal, Canada

URBAN SUSTAINABILITY: SUPPORT FROM ARTIFICIAL INTELLIGENCE AND BIG EARTH DATA
Organized by Daniele Ehrlich, European Commission, Ispra, Italy

SCIENCE ON TAP AAAS Flash Talks

Brief presentations highlighting scientific findings and programs

AI-powered Ultrasound Diagnoses Pneumonia Faster, Better than Experts

Biodiverse Soil Ecosystems: Influencing Climate, Industry, and Life

Building Resilient Communities During Times of High Risk and Uncertainty

The Cancer Paradox: Mutational Secrets Hidden in the Animal Kingdom

Communication, Preparation, and Response: The Effects of Environmental Threats on Health and Behavior

Crop Mutation Application Under Climate Change Mitigates Environmental Pollution

Discipline-based Education Research: Informing a More Effective Undergraduate STEM Experience

The Future of Earth's Ice: A Human Action Story

Gene Editing Goes Global

The Global Disinformation Index: Finding Those Who Corrupt the World's Information

How Decision-Support Technology Pumps Up a Sustainable Groundwater Ecosystem

Iodide: A Primordial Antiperoxidant for Treating Trauma

Let's talk! Bilingualism as One-Health Approach to Understand Neurocognitive and Social Plasticity

Making Medicines Personal: Is It All in Your Genes?

Non-alcoholic Fatty Liver Disease: It's All in Your Head

Population Health Goals for Tomorrow's Earth

Prevalent Plasticizers, Chemotherapeutic Agents, and Disruptions to Embryo Development

Safe Drinking Water for All: Spotting Bacterial Contamination

Stress, Sex, and Inflammation: Metabolic Mediators in the Brain

Science Careers Workshops

Opportunities to gain advice and strategies from experienced STEM professionals

Building A Responsive Network For Tomorrow's Science Communication Needs

Bull's Eye: Developing Specific Aims for a Successful Research Proposal

Careers in State Science, Technology, and Innovation Policy

Citizens, Scientists, and Elections: How Scientists are Engaging in 2020

Communications, Engagement and Advocacy: The Role of Science in Decision-making

Crafting a Narrative for Your Post-academic Career

Envisioning the Global Professional: Uncovering Career Opportunities Worldwide

Envisioning Tomorrow's STEM Workplace: Think Globally, Act Locally

Envisioning Your Career of Tomorrow, Today

European Union Grants: Shake Up Your Research Abroad

Exploring Careers at National Laboratories

Exploring Diverse Avenues to a Career in Science Policy

Friends of the Science Pod: Podcasting, Outreach, and Professional Networking

From the Bench to the Ballot: Scientists in Elected Office

The (Gross) Anatomy of Responding to Peer Review Commentary

How to Make Compelling Outreach Videos When Your Science Seems Dull

How Youth Can Contribute to the Sustainable Development Goals

Impacting the Research Enterprise Through Careers in Research Development

Learning to Manage and Mentor: Skills for Long-term Success in Science

Make 'em Laugh: Science Comedy to Ignite Curiosity and Increase Self-confidence

Navigating Difficult Situations in Public Science Communication

OUT on the Job Search: Finding a Welcoming Environment

Science in the Public Arena: Informing Decision Makers in High-Profile Settings

Shaping STEM Policy without Changing Careers: Local Government Opportunities

Strategies for Securing Philanthropic Funds in Science and Health

Teaching Inclusively with Evidence-Based Strategies

Sponsorship and Exhibit Spaces Still Available

Enhance your visibility at the AAAS Annual Meeting. Don't miss this chance for your research, products, and services to grab the attention of researchers, scientists, engineers, journalists, policymakers, educators, students, hiring managers, and funding agencies.

Visit **aaas.org/meetings** for more information.

Advance registration rates are available now through **January 24, 2020.** On-site registration rates will apply thereafter.

	Advance Rates for AAAS Member for members in good standing	Advance Rates for Non-Member for all other attendees	On-site Rates after 1/24/2020 AAAS Member/Non-Member
General Attendee	$310	$440	$380/480
Postdoc	$135	$260	$135/280
K-12 Teacher	$135	$360	$135/380
Retired Professional	$250	$360	$295/380
Student	$65	$95	$75/105
One-Day	$175	$220	$200/240

REPORTERS: The AAAS Annual Meeting Newsroom will be hosted on EurekAlert! at eurekalert.org/aaasnewsroom

Breath Sampler and Portable Air Supply
Owlstone Medical introduces the ReCIVA-Breath Sampler and CASPER Portable Air Supply, designed to support internal biomarker discovery programs and those of academic and biopharma clients, and to measure levels of known volatile organic compounds (VOCs) on breath. CASPER features a range of benefits, such as uniting the airflow pump and filter into a single casing that connects directly to the ReCIVA Breath Sampler, the use of filter cartridges that can be easily replaced, and active warnings when the cartridges need changing or when temperature or airflow issues occur. In addition, we have implemented new software to monitor subject breathing and VOC collection through ReCIVA, including a simpler user interface and the ability to set different levels of access to settings for different users.
Owlstone Medical
For info: 203-908-4848
www.owlstonemedical.com

Human IgGκ Detection Kit
Sphere Fluidics offers the Cyto-Cellect Human IgGκ Detection Kit, to be used in conjunction with the Cyto-Mine Single Cell Analysis System for rapid, cost-effective measurement of antibody class and productivity in cell line development. The Cyto-Cellect Kit has been developed to provide Cyto-Mine users with a convenient, standardized assay format for the detection of secreted human antibodies in picodroplets. The sensitive, fluorescence resonance energy transfer (FRET)-based assay rapidly detects the presence of human immunoglobulin G (IgG) with kappa (κ) light chain after secretion from encapsulated cells, to characterize and measure antibody production. Rapid detection of IgG enables high-throughput identification and selection of cells with the highest productivity for more efficient cell line development. With optimized ratios of Sphere Fluidics' Human IgGκ Detection probes, Cyto-Cellect offers a sensitive, robust assay ideal for streamlining cell isolation and titer determination. This animal origin–free reagent kit delivers confidence in the reproducibility and consistency of the assay and conforms to regulatory requirements in cell line development.
Sphere Fluidics
For info: 888-258-0226
spherefluidics.com/cyto-cellect-human-igg?-detection-kit

Culture Media
FUJIFILM Irvine Scientific introduces the BalanCD Gal Supplement. This formula has been developed to increase galactosylation in biotherapeutic development, and to help achieve desired glycan profiles for improved product quality and antibody efficacy. Glycosylation is a critical aspect in development and optimization of protein-based therapeutics. The type and level of glycosylation can significantly affect antibody binding, function, and therapeutic efficacy. Galactosylation (the glycosylation of galactose) is a key method for controlling product quality due to its significant impact on antibody function, and the ability to control galactosylation levels through the cell culture medium. BalanCD Gal Supplement is a chemically defined, animal component–free formula designed to increase N-linked galactosylation with a scalable protocol, offering researchers a simple method to control the process through the cell culture medium. The supplement is effective at low concentrations and is compatible with any basal growth medium.
FUJIFILM Irvine Scientific
For info: 800-577-6097
www.irvinesci.com/industrial-cell-culture

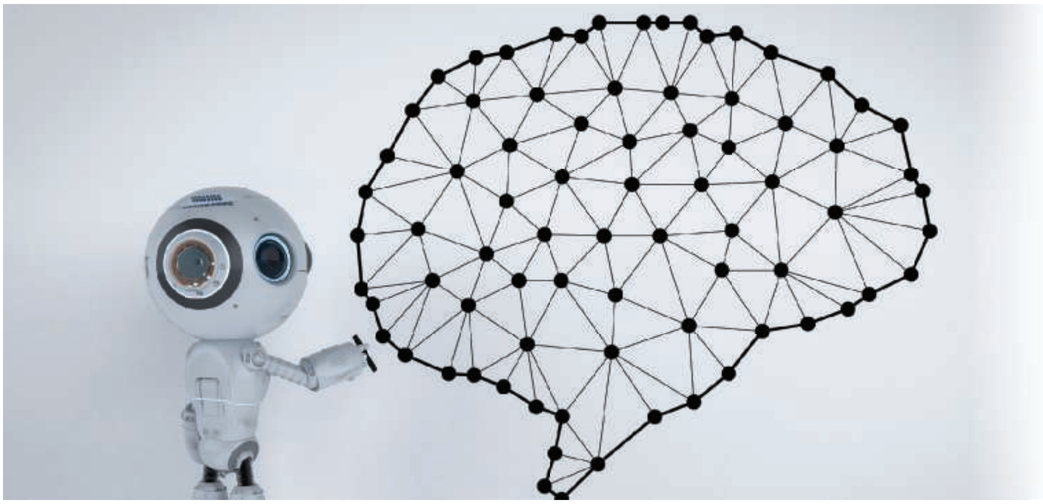
EV Targeting Kit
The XStamp Pro Streptavidin EV Targeting Kit gives you virtually unlimited target specificity options. Just decorate the outside of your extracellular vesicles (EVs) with XStamp Pro Streptavidin and then attach the biotinylated targeting moiety partner of your choice, such as a biotinylated single-chain variable fragment (scFv) antibody. Save time and simplify your EV engineering workflow with XStamp Pro. While conventional methods for creating target-specific EVs are highly robust, they involve cloning and manipulation of the producer cells, which can be challenging, especially if the cells are hard to transfect. To streamline this process and help researchers more quickly and easily create EVs that preferentially localize to a specific target, System Biosciences turned to an exciting new technology, XStamp Pro, which works on already purified EVs—no cloning or transfection required. Simply incubate purified EVs with the XStamp Pro reagent for 1 h and then isolate your target-specific EVs.
System Biosciences
For info: 888-266-5066
systembio.com

Confocal Objectives
Molecular Devices announces water immersion objectives as a high-performance customization for the ImageXpress Micro Confocal High-Content Imaging System. Water immersion objectives can offer signal increases of up to fourfold, which can help scientists see deeper into 3D and thick tissue samples at lower exposure times. These objectives also improve *z*-resolution and spherical aberrations; enable sharper, crisper images; and offer access to more data than ever before. The Micro Confocal system is a high-content solution that can switch between widefield and confocal imaging of fixed and live cells. It can capture high-quality images of whole organisms, thick tissues, 2D and 3D models, and cellular or intracellular events. Molecular Devices can successfully tailor the system to include customized software and hardware as well as integration of other lab components, such as incubators, liquid handlers, and robotics.
Molecular Devices
For info: 800-635-5577
www.moleculardevices.com

Miniature Solenoid Valve
The genvi solenoid valve provides outstanding application versatility and unmatched reliability at an economical price. One of the key objectives in developing the genvi solenoid valve was to challenge the limitations of current fluidic technologies by allowing designers to use a miniature, low-power solenoid valve without compromising system-level flow-rate requirements. As a result, the valve is capable of achieving minimum flow rates of 30 to 40 standard liters per minute (SLPM) at 15 psi differential (psid), and its spike-and-hold profile results in ultralow power consumption (318 mW) during continuous duty operation on air and compatible gas applications. These valves are designed to provide response times of less than 10 ms in applications with an operating temperature range from 40°F to 120°F. Typical applications for the valves include molecular diagnostics, oxygen delivery, environmental monitors, compression therapy, and breath analyzers.
Lee Products
For info: +44-(0)-1753-886664
www.leeproducts.co.uk

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.



Science 2019 TOP EMPLOYER

PHOTO: © LIGHTSPRING/SHUTTERSTOCK.COM

Top employers: Breakthroughs, impact, and purpose

The 17th annual Top Employers Survey features a surprise: Alnylam Pharmaceuticals, an RNA-interference therapeutics company headquartered in Cambridge, Massachusetts, earned the No. 1 spot in its first appearance in the survey rankings. Some responses from the biotechnology and pharmaceutical industry are similar to those in previous surveys, however. Respondents valued innovation above all, while noting industry changes around drug pricing, regulations, and policies as well as an increasing emphasis on artificial intelligence and machine learning. **By Chris Tachibana**

Followers of the annual Top Employers Survey from *Science* Careers will notice something new this year. The highest ratings in 2019 went to newcomer **Alnylam Pharmaceuticals**. The U.S.-based company of more than 1,200 employees develops RNA-interference (RNAi) therapies. In the three previous years, the top employer was **Regeneron Pharmaceuticals** in New York, which is No. 2 this year, followed by the Delaware-based pharmaceutical company **Incyte.** "We're very excited," says Alnylam CEO John Maraganore regarding the company's Top Employer status. "We've grown a lot lately and our success depends on having a highly engaged team."

Many other features of the survey remain unchanged, however. As happened in recent years, more than 7,500 people responded. About 95% reported working in the biotech and pharma industry, and 80% were age 30 years or older. This year, the proportion of survey respondents from North America increased to 72% from 63% in 2018. The proportion from Europe dropped from 24% to 19%, and the fraction from the Asia/Pacific Rim fell from 9% to 7%.

Innovation has been a leading driver of top employer status since the survey began in 2002, and this year was no exception. Other reasons for recognition as a top employer were treating employees with respect and having company values that align with theirs. Being socially responsible and having leadership that can make needed changes were also important characteristics of top companies.

The most noteworthy current and expected changes in the industry, based on open-ended comments from survey respondents, focused on cost-cutting, pressures around drug pricing, and the political environment, including drug regulation policies and changes at the U.S. Food and Drug Administration (FDA). Mergers and acquisitions, outsourcing, and the impact of artificial intelligence (AI) were noted, along with the rising pace of industry research.

Representatives of some top companies gave their perspectives on these and other issues. They discussed how their organization maintains an innovative edge and excels as a workplace, and how the increasing use of AI affects work and work culture.

The top five innovators: Breakthroughs in products and pricing

Alnylam joined the Top Employers list with a splash, reaching the top spot in its first year of inclusion in the survey. It's been a year of breakthroughs for the company. In August 2018, Alnylam received the FDA's first-ever approval for an RNAi therapy. Alnylam's Onpattro treats neurological symptoms, such as numbness in patients with hereditary transthyretin-mediated amyloidosis. This rare, potentially fatal disease affects about 50,000 people worldwide.

Maraganore says that having the first RNAi therapy on the market is a clear sign that Alnylam is an innovative leader, which is what survey respondents valued most in a biotech or pharma company. "We are pioneers in bringing a whole new class of medicines to market," he says. "And before that, we pioneered bringing this technology to clinical trials."

Alnylam regularly conducts internal work culture surveys and develops improvement plans based on the results, Maraganore says: "We view feedback from employees as a big gift, and we harness that to continuously make our company better." Employees notice innovations in benefits, Maraganore says, for example, an in-house diversity and inclusion team that works on issues such as equity in race, gender identity, and sexual orientation. This initiative aligns with the

cont.>

employee value of social responsibility, consistently rated in the survey as a characteristic of the best companies. Diversity is also a trending issue in the industry. Maraganore just stepped down after two years chairing the Biotechnology Innovation Organization (BIO) Board of Directors. At the last BIO International meeting, BIO launched the Right Mix Matters campaign to provide companies with resources to increase diversity in leadership positions.

Maraganore also notes employee work–life benefits, such as programs for working at home and for spending time on an exploratory project not directly related to Alnylam R&D. This opportunity can pay off for the company in a big way. "We recently figured out how to deliver our drugs to the central nervous system, which opens up our pipeline to a range of neurodegenerative diseases," he says. "That opportunity happened because we gave a small group of employees the freedom to take 20% of their time to explore a new idea."

Drug pricing is at the nexus of a number of topics—including access to medicines, cost-cutting, and politics—that survey takers raised when asked to name notable industry changes. Many companies with groundbreaking but high-cost drugs are developing novel pricing schemes. Maraganore mentions Onpattro pricing related to both innovation and social responsibility, saying that company representatives proactively met with health care payers (insurance companies and other health plan sponsors) to negotiate value-based reimbursement. In these plans, payments made by payers to the company are linked to patient response to drug therapy. "With other drugs," he says, "you pay even if they don't work. We believe in our product so we're willing to put skin in the game with value-based reimbursement."

At No. 4, **Merck KGaA** is a contrast to the newer companies in the top five: Merck KGaA celebrated its 350-year anniversary in 2018. (The company is legally independent from U.S.-based Merck & Company and has headquarters in Darmstadt, Germany.) Nonetheless, in common with other top employers, Merck KGaA prioritizes thinking ahead, adapting, and communicating. This strategy is how the company takes advantage of opportunities and innovations, said member of the Executive Board and CEO of Healthcare Belén Garijo. In an email, she said Merck KGaA's long-term success centers around maintaining connections to all stakeholders, including employees, business leaders, and customers.

Top Twenty Employers

2019 Rank	2018 Rank	Employer (global headquarters)	Innovative leader in the industry	Treats employees with respect	Is socially responsible	Work culture values aligned	Leadership makes changes needed
1	–	Alnylam Pharmaceuticals (Cambridge, MA)	•	•		•	
2	1	Regeneron (Tarrytown, NY)	•		•	•	
3	2	Incyte (Wilmington, DE)		•	•	•	
4	5	Merck KGaA (Darmstadt, Germany)		•	•	•	
5	–	Spark Therapeutics (Philadelphia, PA)	•	•		•	
6	7	Biocon Limited (Bangalore, India)	•		•		•
7	3	Novozymes (Copenhagen, Denmark)	•	•	•		
8	9	Genentech (South San Francisco, CA)	•	•		•	
9	16	Eli Lilly and Company (Indianapolis, IN)		•	•	•	
10	14	Syngenta (Basel, Switzerland)	•	•	•		
11	4	Moderna (Cambridge, MA)	•	•	•		
12	6	Vertex Pharmaceuticals (Boston, MA)	•		•		•
13	10	AbbVie (North Chicago, IL)			•	•	•
14	12	Roche – excluding Genentech (Basel, Switzerland)	•		•		•
15	11	AstraZeneca (Cambridge, UK)		•	•	•	
16	–	GlaxoSmithKline (GSK) (London, United Kingdom)		•	•	•	
17	18	Abbott (Abbott Park, IL)		•	•		•
18	–	Pfizer (New York, NY)			•	•	•
19	13	Novartis (Basel, Switzerland)	•	•		•	
20	8	Novo Nordisk (Bagsværd, Denmark)		•	•	•	

The 20 companies with the best reputations as employers and the top three driving characteristics for each company, according to respondents in the 2019 survey undertaken for the *Science*/AAAS Custom Publishing Office.
The companies without a 2018 rank did not receive enough mentions to qualify or did not receive a high enough ranking from the 2018 survey.
The fifth attribute in the above table is "Has top leadership that successfully makes changes needed to keep the organization moving in the right direction."

Spark Therapeutics, like Alnylam, celebrates its first year of survey inclusion by entering the survey at No. 5. Like Alnylam, Spark brought a groundbreaking therapy to market in 2018 with the first U.S. commercial sales for a gene therapy product. Luxturna treats vision loss from a rare, inherited retinal dystrophy disease.

Founded in 2013, Spark is headquartered in Philadelphia, Pennsylvania, and has more than 400 employees. Katherine High, president and head of R&D, says the company's innovation is demonstrated by Luxturna and four other gene therapies in clinical trials, with more in the pipeline. High recently talked to an employee with a background in more traditional pharmaceutical work who noted the transformative effect of gene therapy. "A lot of drug programs are just trying to have a narrow margin of superiority over others," High says. "With gene therapy, as long as the program is well thought out, we see very clear therapeutic effects."

Spark was highly rated by survey participants for having a work culture that aligns with employee values. To describe the company's culture, High uses the adjective "dynamic." One employee, she says, noted that "there's high speed, there's warp speed, and there's 'Spark speed.'" Employees see their programs progressing, High explains, and watch their hard work move products from preclinical to clinical stages—and in Luxturna's case, to commercial success. "All that is exciting," she says.

On the topic of drug pricing, High lists factors that her company considers when setting prices. The diseases for which Spark is developing gene therapies, such as retinal dystrophy, have no available treatments or have high unmet needs. An example of the latter is hemophilia, which requires frequent infusions of clotting factor. A one-time gene therapy intervention could save money over treatments that cont.>

must be administered over a lifetime and may treat only symptoms, not causes. Pricing also needs to reflect the investments of developing a one-time therapy and should allow a company to be sustainable, High notes.

To provide patients access to needed therapies, Spark is pursuing several strategies in the United States focused on health care payers, High says. The strategies include installment plans for payments as well as outcome-based rebates derived from the same principles as Alnylam's value-based reimbursements. For example, if patients don't achieve expected outcomes based on Phase III clinical trial results, payers are eligible for a rebate from Spark. Because gene therapy is administered at only a few medical centers in the country that have specially trained personnel, agreements with payers ensure that no matter where patients are treated, they pay in-network rates, as though they received care in their home area.

In another novel approach called "buy and bill," payers rather than medical centers purchase the therapy. This plan reduces the financial risks to medical centers—for example, from buying therapies that patients don't end up using—and eliminates markup costs that payers might otherwise incur. Spark is also looking into installment plans for payments tied to therapy effectiveness, but because of U.S. health care complexities, this requires state-by-state arrangements.

With regard to survey participants' comments about mergers and acquisitions, High has direct experience with employee concerns about this situation. Roche is in the process of buying Spark after the boards of both companies unanimously agreed to the acquisition. Employees are positive about the merger, High says. "We'll have to see how it unfolds over time, but we see this as an opportunity to access additional financial resources to push our work forward. I personally think it's exciting," she adds. "Roche is a world-class drug developer."

DRIVING CHARACTERISTICS OF TOP EMPLOYERS

2019	2018
1. Innovative leader in the industry	1. Innovative leader in the industry
2. Treats employees with respect	2. Treats employees with respect
3. Is socially responsible	3. Is socially responsible
4. Has work culture values that are aligned with personal values	4. Has work culture values that are aligned with personal values
5. Has top leadership that successfully makes changes needed to keep the organization moving in the right direction	5. Has top leadership that successfully makes changes needed to keep the organization moving in the right direction
	6. Has a clear vision of where it is headed

* Driving characteristics are listed in descending order of impact on overall employer rankings.
** Shaded background indicates the characteristics in common for the two years.

IMAGE: © SPAINTER_VFX/SHUTTERSTOCK.COM

Emphasizing persistence and purpose

The acquisition would add Spark to the Roche group, which includes **Genentech,** this year at No. 8 in the survey. Founded in 1976, with headquarters in South San Francisco, the well-established Genentech is a contrast to Spark. While this is Spark's first year in the survey, Genentech is the only company to have been one of the top employers since the survey started in 2002, setting the standard for the entire field.

While the term "speed" (along with "well thought out") comes up in High's description of the Spark culture, Mike Varney, Genentech's executive vice president of research and early development, uses the word "patience" to describe his company's scientific approach. The Roche group has room for both Spark and Genentech, however, because of a common overall culture. "We all value rigorous science," Varney says, "but Genentech and others in the Roche group maintain their own subculture. We build the organization the way we want and create the kind of work culture that will facilitate our innovation."

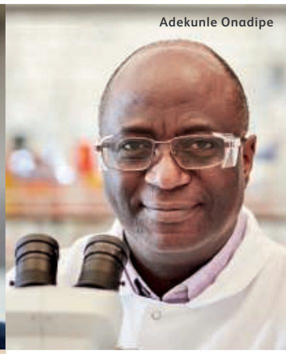
Genentech maintains its status as an innovative leader by harnessing technology to deeply understand the biology of disease, Varney says. Initially, the company's founders used this principle to translate the technology of gene cloning into medicines such as hormones. Later, company researchers focused on converting the specificity and binding affinity of antibodies into immuno-oncology therapies.

"This business requires understanding complex biology," Varney says, "so patience is a virtue and persistence is a huge component of success. We're willing to put the time and resources into solving problems." He notes that distinguishing features of the company are its high ratio of discovery biologists to other employees and the commitment of its researchers to use data to guide their projects.

Varney agrees with survey participants about the increasing pace required to move products through pipelines. "There's no question that time matters," he says, but emphasizes that the company's approach is efficient and strategic in the end. Researchers don't waste time pursuing medicines that their data do not support, but when their findings show promise, they continue. This is one reason so many of Genentech's products are first-in-class therapies, Varney observes. "We believe in our biology so strongly that we stay in the game when others drop out." An example, he says, is that company persistence resulted in their AKT (protein kinase B) inhibitor, now in clinical trials for cancer therapy.

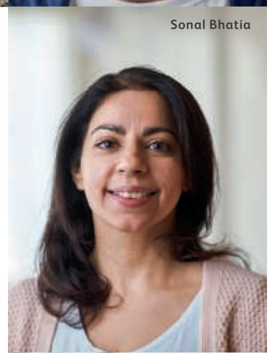
One of Genentech's strengths in the survey was quality research with talented employees. An experiment-focused company attracts action-oriented people interested in exploring the unknown, Varney says. "There's no innovation without experimentation. In an innovative environment, you take action," he says. "In a noninnovative environment, you analyze." Genentech has long had programs that encourage researchers to explore avenues not always directly related to their company work. Currently, research leaders can apply for internal innovation funds or for a postdoc to work on an industry or academic project.

In the midst of a general emphasis on multidisciplinary teams and flat structures, Genentech also holds to a single decision-maker model for its research teams. "You can have a freewheeling team with freedom to explore," Varney says, "but someone has to make the decisions and point the team members in the same direction." The decision-maker's goal is to choose based on data and input from the team. This model, Varney says, "provides organizational clarity." cont.>


Iman Jilani

Adekunle Onadipe

Guoyun Bai

Ricky Fernandes

Herbert Medina

Sonal Bhatia

Regina McDonald

Charles Cain

Farhan Hameed

Karen Walters

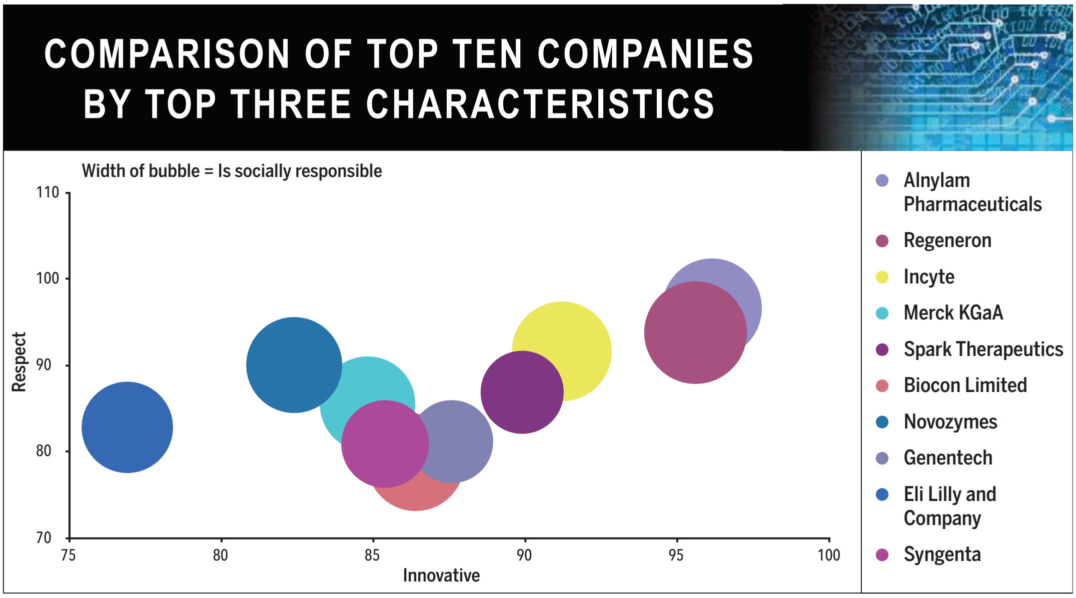

Elaine Ravasco

Adekola Alagbe

Kelly Voight

Patrick McCann

At No. 9 this year was **Eli Lilly,** moving up from No. 16 in 2018. The company is headquartered in Indianapolis, Indiana, but has a presence around the world. In an email, Terri Grant, vice president for human resources at Lilly Research Laboratories and Lilly Oncology, noted several factors that are attractive to employees. These include a focus on transformational medicines and a collaborative approach involving scientists, physicians, academic researchers, and contributors from health care organizations. Like Alnylam, Lilly is also responding to employee feedback by working to increase diversity, including in management and leadership positions.



Comparison of the top 10 companies on the basis of 3 of the top drivers (scored out of 100): *Is an innovative leader in the industry* (x-axis), *Treats employees with respect* (y-axis), and *Is socially responsible* (bubble width).

IMAGE: © TITIMA ONGKANTONG/SHUTTERSTOCK.COM

A strength of Lilly compared to companies with similar rankings in the survey is social responsibility. Grant points to two programs that make the company stand out in this area. On Lilly's annual Global Day of Service, employees volunteer in their communities. Also, the company's Connecting Hearts Abroad program sends about 100 employees a year on a two-week volunteer service assignment, for example to South Africa or Mexico. Participants help with health-related issues ranging from diabetes care to mobile community health screenings, gaining inspiration and experience that enhance their sense of purpose as employees.

Syngenta, an agriculture company headquartered in Basel, Switzerland, was No. 10 this year, moving up from No. 14 in 2018. Gusui Wu, head of seeds research, has been with the company for only a year, but says a reason for the rise in employer standing might be that after a period of change (including acquisition by ChemChina in 2017), Syngenta now has a clearer vision for the future. Both leadership and employees see a more definable role for Syngenta in the industry and in how it contributes to society. "Our industry is historically not seen as sustainable, but we have a vital role in food security," Wu says, "so there's now more emphasis on agricultural sustainability and meeting the challenge of climate change."

In fact, a strength of Syngenta relative to companies with similar rankings is corporate responsibility. Wu observes that company goals of helping feed the world while protecting the planet through sustainable agriculture align with employees' personal values. By working on products that help farmers address issues such as drought, changes in crop pests, and diseases that arise from the climate crisis, he says, "our scientists feel they're doing good by coming to work every day."

According to Wu, Syngenta is in a unique position to help developing economies by lifting their agricultural productivity—for example, with products for insect and disease control. Syngenta research sites, corporate offices, and production plants are held to sustainability and environmental health standards with periodic audits, which may be another responsible action noticed by employees.

Syngenta has made recent outreach efforts to the public and especially customers. "We believe people have misperceptions about what we do," Wu says. To be clearer about the benefits of agricultural technology, this year Syngenta conducted a 90-day initiative consisting of listening sessions with, for example, consumers and farmer groups. One result was a long-term partnership with The Nature Conservancy, which got positive feedback from employees. Another initiative, focusing on soil health, aims to develop technologies that reduce nutrient loss in soil.

Earlier this year, Syngenta began a project to facilitate R&D collaborations with farms, which has connected more than 100 Syngenta scientists with large agricultural operations, Wu says. The project's goals are sharing data, demonstrating technology, and collecting information on product performance. "Farmers get to see the technology we're working on," he explains, "and scientists get direct input and feedback from farmers who will be using our products." Also, like Alnylam, Genentech, and other companies, Syngenta has a competitive internal funding program for researchers to explore high-risk, high-return, creative projects separate from their product-development work.

Since agrobusiness is going through a period of consolidation, Wu has insights about employee management during mergers and acquisitions. The 2017 acquisition of Syngenta by ChemChina was for the purpose of expansion and growth, and to help China increase agricultural productivity, he says, not for cost-cutting and reductions. Still, when two companies come together, regardless of the reason, processes, procedures, and cultures will change. Recognizing the inevitable disruption that Syngenta faced, the company instituted change management programs so employees understood the business rationale for the acquisition and what it meant to them. "The worst situation is leaving employees and the organization in uncertainty," Wu says. "The ambiguity can be unsettling."

Communication is critical; it's especially important to ensure that employees hear frequently from corporate leaders and their own supervisors, Wu says. "Even if we don't

cont.>

Featured participants

Alnylam Pharmaceuticals www.alnylam.com	**Merck KGaA** www.merckgroup.com/en
Eli Lilly www.lilly.com	**Regeneron** www.regeneron.com
Genentech www.gene.com	**Spark Therapeutics** sparktx.com
GSK www.gsk.com	**Syngenta** www.syngenta.com
Incyte www.incyte.com	

have answers, we explain that we're trying to get clarity on questions. We report progress and keep employees as informed as possible."

Bringing in AI and ML

In addition to pressures related to drug pricing and mergers and acquisitions, survey participants consistently note industry changes, including the rising use of automation, AI, and machine learning (ML) in research. At Alnylam, for example, Maraganore says AI and ML are enhancing effectiveness in multiple ways, from identifying sequences for designing RNAi drugs to locating patients who might benefit from Onpattro.

At Merck KGaA, Garijo says AI tools are expected to increase efficiency and effectiveness. The company has several AI-related agreements, including with AI drug-design company Iktos and proteome-screening company Cyclica. Merck KGaA also received a U.S. patent for a system that uses AI to protect supply-chain integrity. The system links physical objects, such as equipment parts or pills, with digital signatures that are securely stored with blockchain technology. This process ensures the authenticity of medicines but could also be used for products such as food and electronic devices, Garijo explains.

AI tools fit the Genentech commitment to understanding the biology behind diseases, says Varney. He sees drug discovery as a sorting exercise, narrowing targets and candidates from a large pool down to the most promising. The company integrates data scientists within research teams to apply AI where it can make sorting more efficient. One example is an application that rejects small-molecule drug candidates when data indicate they are likely to be quickly metabolized. Another applies data to identify tumor-associated proteins that will be antigenic and easily displayed to the immune system for personalized cancer vaccines. By making sorting more efficient, "AI frees up scientists' time so they can think and be creative and do other work that machines can't do," Varney says.

GSK (GlaxoSmithKline), headquartered in Brentford, United Kingdom and at No. 16 this year, has long been a leader in AI and ML. GSK has used AI for traditional R&D, such as small-molecule drug discovery. Recently, the global pharmaceutical company stepped up its AI–ML game.

In 2018, Hal Barron became GSK's chief scientific officer and president of R&D. He is directing new investments in R&D, particularly in immunology, human genetics, and advanced technologies. GSK is focusing its technology development on the intersection of human genetics data, functional genomics, and AI and ML to help understand human disease on a cellular level. Barron explains two points driving GSK R&D: (1) Less than 10% of drugs that enter clinical testing go to market, and

DEMOGRAPHICS

GENDER:
50% Male, 44% Female, 6% No response

EXPERIENCE:
71% have 10 or more years work experience

HIGHEST DEGREE EARNED:
33% Doctorate, 31% Master's, 30% Bachelor's, 6% Other

COMPANY TYPE:
21% Biotech, 40% Biopharma, 34% Pharma, 2% University, 3% Other
More than 9 out of 10 work in private industry.

NATURE OF WORK:
13% Basic Research, 17% Applied Research, 26% Development, 8% Production, 11% QA/QC or Regulatory Affairs, 9% Executive, 16% Other

GEOGRAPHY:
72% North America, 19% Europe, 7% Asia/Pacific Rim, 2% rest of world

(2) genetic validation increases the likelihood a medicine will succeed. This is why GSK developed partnerships with human genetics organizations such as 23andMe and Open Targets to use data to help identify new drug targets.

To validate targets, GSK plans to incorporate CRISPR gene-editing technology for functional genomics through a partnership with the University of California. Researchers will use CRISPR to test how altering candidate genes or their expression affects human cells in vitro. Functional genomics generates "trillions of datapoints," Barron says, so ML is essential for the next step—integrating the data and helping to understand relationships between genes and how mutations relate to disease. To support this work, GSK is building infrastructure, including an in-house data science group, a platform for integrating large datasets, and automated tools for data analysis. Barron expects these innovations to increase R&D speed from discovery to clinical trials to market.

The bottom line: Having an impact

The addition of AI and ML to drug development and marketing is speeding breakthroughs across the industry. More mergers, acquisitions, and consolidations are expected by survey participants, who also noted Brexit and upcoming U.S. elections as changes expected to affect the industry.

The bottom line for pharma and biotech companies, however, is that employees need to feel that their company supports them in doing high-quality, rigorous, impactful work, Maraganore notes. "I believe our employees are invigorated to work for a company that is bringing innovative medicine to patients," he says. "They're generating something brand-new and transformative, which creates a sense of pride and a sense of purpose."

To read this article online, view top employer profiles, and listen to podcast interviews, visit: ScienceCareers.org/TopEmployers

Chris Tachibana is a freelance writer who specializes in life sciences.

TOP EMPLOYERS

POSITIONS OPEN

UNIVERSITY of ROCHESTER MEDICAL CENTER

Assistant Professor/ Associate Professor
Center for Biomedical Informatics (CBI) and Department of Microbiology and Immunology

The University of Rochester Medical Center (URMC) is expanding its research activities in the area of bioinformatics. We are seeking investigators at the Assistant Professor or Associate Professor level with independent and collaborative research programs. Six to ten faculty will be recruited into tenure track positions over the next three years. These faculty will join an existing group of collaborative, multidisciplinary colleagues performing research in complimentary fields.

We are searching for investigators with expertise in the broad areas of bioinformatics, computational, and systems biology as applied to a research program studying infectious disease and immunity. Successful candidates will hold appointments in the CBI and the Department of Microbiology and Immunology. Investigators with research experience in the following areas would be especially qualified for this position: Genome and transcriptome informatics, systems biology, population genomics, machine learning, data mining, computational modeling, multi-dimensional integration of clinical, genomic, microbiome, metabolome, gene expression and epigenetic data.

URMC offers attractive start-up packages and has a strong commitment to career development. The successful candidate is expected to develop a competitive research program, attract external funding, and participate in graduate education. Major recent institutional investments have created an outstanding research data–driven infrastructure, exemplified by the Health Sciences Center for Computational Innovation, and the recent $50 million University investment to build the Goergen Institute for Data Science.

Interested individuals should submit a cover letter, CV, statement of research interests/plans, and arrange to have three letters of recommendation sent to: **anne_reed@urmc.rochester.edu**. Inquiries can be directed to Dirk Bohmann (dirk_bohmann@urmc.rochester.edu) or David Topham (david_topham@urmc.rochester.edu).

The University of Rochester is an Equal Opportunity Employer and has a strong commitment to diversity and actively encourages applications from candidates from groups underrepresented in higher education.

UR MEDICINE | WILMOT CANCER INSTITUTE

Aging/Cancer Biology Faculty Position

The Wilmot Cancer Institute (WCI) at the University of Rochester Medical Center is currently recruiting tenure-track faculty as part of a major expansion of its translational and basic science research base. The biology of aging is an area of expanding interest at the WCI and the University as a whole. The successful candidate will join a growing multidisciplinary research community with an overarching focus on aging and its complex interaction with cancer. Scientists studying aging in relation to cancer cell metabolism, cancer stem cell biology, RNA biology, DNA repair, cancer (epi) genomics and systems biology, microenvironment influences on tumor behavior/immunology, and therapeutic resistance among other topics would be well suited for this position. The Wilmot Cancer Institute is the hub of cancer research at the University of Rochester, attracting more than $20M in cancer-directed funding annually. Formal research programs exist in Cancer Biology, Tumor Microenvironment and Immunology, and Cancer Control and Survivorship. The Wilmot Cancer Institute and UR Medicine are the major providers of comprehensive, multidisciplinary cancer care in central New York State and the Finger Lakes region, with >6000 new patients every year and an extensive clinical trials program providing opportunity for innovative translational research.

Candidates should hold a PhD and/or MD degree and have a demonstrated track-record of research accomplishment in an area relevant to aging and cancer biology. Scientists making use of genomic/epigenomic approaches, bioinformatics, model organisms, tumor model systems and/or human specimens to address problems of translational relevance are strongly encouraged. New faculty will benefit from vibrant graduate/professional training programs and state of the art infrastructure and core facilities, as well as a strong Institutional commitment to career development. Appointments will be made at the Assistant Professor level although outstanding candidates at other levels will be considered, with commensurate expectations of research and funding accomplishment. Departmental affiliation will be determined according to best fit. Interested Individuals should submit a CV, statement of research interests/plans, pdfs of two key publications, and three letters of recommendation to the Search committee chairs **Vera Gorbunova, PhD and Dirk Bohmann, PhD** c/o Jaycee Bristol at **Jaycee_Bristol@urmc.rochester.edu**. Review of candidates will start **December 1, 2019**.

The University of Rochester is an Equal Opportunity Employer and has a strong commitment to diversity and actively encourages applications from candidates from groups underrepresented in higher education.

Jiangsu, referred to as "Su", exists as one of the most open, prosperous and innovative regions in the Yangtze River Delta economic circle of China. With its economy thrives, culture and education also boom.

Here, Jiangsu boasts picturesque scenery. There are refined ancient towns, famous temples of several thousand years, classical gardens, beautiful lakes and mountains, emperors' tombs and historical capital city sites.

Here, Jiangsu enjoys a fame for celebrities. In the past, there were Zu Chongzhi, Shen Kuo, Xu Guangqi, Xu Xiake, etc. Today, there are Hua Luogeng, Zhou Peiyuan, Mao Yisheng and Qian Weichang.

Here, Jiangsu prospers in higher education. There are 142 well-known universities such as Nanjing University, Southeast University, Jiangsu University, Suzhou University and China Pharmaceutical University. There are also other 8 adult colleges and 25 independent colleges.

Today, the innovation capacity in talent training, scientific research, social service and cultural inheritance in Jiangsu universities has been continuously improved. And the schools always aim to be an important base for cultivating all kinds of high-quality talents, to become an important force for knowledge discovery and technological innovation. They strive for advanced thinking and excellence in higher education. It cannot be neglected that the schools are playing a major role in supporting innovation-driven development strategies, serving economic and social development, and promoting China's outstanding traditional culture.

Also, Jiangsu colleges and universities positively build talented highlands. To persist in the idea of"introducing high-end talents, achieving full-time engagement, innovating the hiring mechanism and creating excellent environment", the Jiangsu Specialized Professor Program will be continuously implemented. It will give priority to recruiting talents at home and abroad who have great influence in the international academic community in the field, to accelerate the collection of leading talents and academic innovation teams.

In the future, more universities and disciplines in Jiangsu will enter the world's first-class ranking, further release the competitiveness and influence of Jiangsu's higher education.

We welcome excellent scholars interested in applying for talent programs to contact us through AcaBridge (**consultant@acabridge.edu.cn**), which provides one-on-one consultations.

中国药科大学
CHINA PHARMACEUTICAL UNIVERSITY

Brief Introduction of China Pharmaceutical University

China Pharmaceutical University (CPU), situated in the historical and cultural city of Nanjing, is one of the "211 project" key universities affiliated to the Ministry of Education of China. It has been known for its long history and its leadership role in China's pharmaceutical academic field. The University was founded in 1936 as China's first independent four-year National College of Pharmacy.

The University provides a wide variety of disciplines including Sciences, Medicine, Engineering, Economics, Management, Arts and Law. The first-level discipline of Pharmacy was enlisted as National Key Discipline. In 2018's ESI statistics, our pharmacology and toxicology ranked the world's top 1‰, chemistry, clinical pharmacy and Biology & Biochemistry programs ranked the world's top 1%.

Faculty Positions at China Pharmaceutical University: Nanjing, China

We Are Hunting for

a) **High-Level Talents** selected in one of the national talent plans, including:
- "Chang Jiang Scholars Program of Ministry of Education of China"
"Distinguished Young Scholars Program of the National Natural Science Foundation of China"

b) **Young Talents** selected in one of the national talent plans, including:
- "Chang Jiang Young Scholars Program of Ministry of Education of China"
- "Excellent Young Scholars of the Ten Thousand Plan"
- "Excellent Young Scholars Program of the National Natural Science Foundation of China"

c) **Specially-Appointed Professors or Associated Professors**

Requirements:

(1) Have gained PhD degree from world top University/Institute and under 40/35 yrs old. (2)Good publication record in leading journals

Areas

a) **Life sciences:** Neuroscience, Microbiology, Botany, Biophysics, Biochemistry and Molecular Biology, Genetics and Bioinformatics, Cell Biology, Immunology, Biomechanics and Tissue Engineering, Physiology and Integrated Biology, Synthetic Biology, Food Science

b) **Medical Science:** Respiratory System, Circulatory System, Digestive System, Urinary System, Endocrine System/Metabolism and Nutrition Support, Blood System, Nervous System and Mental Diseases, Pharmacology, Traditional Chinese Medicine

c) **Information Science:** Medical Information Monitoring and Processing, Big Data Processing

d) **Other related disciplines**

What We Offer

a) **High-Level Talents:** Salary no less than 1,000,000 RMB per year; Housing/settling allowance 5,000,000 RMB; Start-up funds 10,000,000 RMB; The laboratory space no less than 300m2..

b) **Young Talents:** Salary no less than 500,000 RMB per year; Housing/settling allowance 4,000,000 RMB; Start-up funds 4,000,000 RMB; The laboratory space no less than 150m2.

c) **Specially-appointed Professors:** Salary 400,000 RMB per year; Housing/settling allowance 500,000 RMB; Start-up funds 2,000,000 RMB.

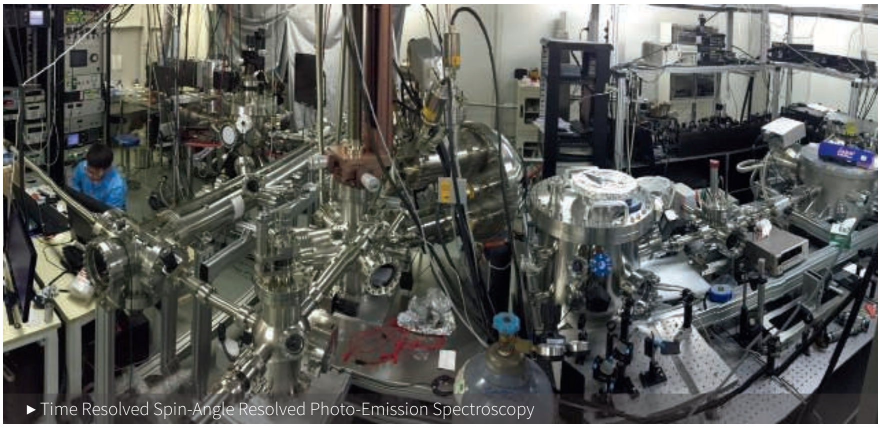
Specially-appointed Associated Professors: Salary 290,000 RMB per year; Housing/settling allowance 300,000 RMB; Start-up funds 1,000,000 RMB.

How to Apply:

a) CV with degree certifications attached;
b) A research plan for the next 5~10 years (no more than 1000 words);
PLEASE TITLE YOUR EMAIL AS "Your Name+GCC+Your Research Areas".

Contact Us

Application materials mentioned above should be sent to:
Personnel Department, China Pharmaceutical University
Contacts: Fan Wang
Tel: +86-25-86185090
Email: rcb@cpu.edu.cn
Address: #639 Longmian Avenue, Nanjing, 211198, P.R.China.
Home page of China Pharmaceutical University: www.cpu.edu.cn



▸ Time Resolved Spin-Angle Resolved Photo-Emission Spectroscopy

The School of Electronic Science and Engineering (ESE) in Nanjing University, established in 2009, can be traced back to the radio-physics program and semiconductor physics program from the 1950s. It is a truly inter-disciplinary program covering the research topics spanning from semiconductor physics, superconductor electronics, advanced devices, integrated circuits, to system applications. From the beginning, the mission of the School is to innovate information technologies through first-class research and education.

▸ Micro Fabrication and Integration Technology Center

History and Development

In 2009, the School of ESE was established as part of the Nanjing University's initiative to strengthen engineering research. It currently hosts 175 faculties, including 55 professors and 44 associate professors. Among them are 2 academicians of Chinese Academy of Sciences, 3 academicians of Chinese Academy of Engineering, 1 IEEE Fellow, 2 chief scientists of the national 973 Program, 8 distinguished Changjiang professors and 8 recipients of the NSFC Distinguished Young Scientists. The School always tries its best effort to offer a most favorable environment for faculty development to construct a first-class team of faculty that facilitates its sustainable development.

The School has established more than 15 national and provincial research platforms, including Jiangsu Provincial Key Laboratory of Functional Materials for Information Optoelectronics, Jiangsu Provincial Key Laboratory of Advanced Manipulating Techniques of Electromagnetic Waves, Jiangsu Provincial Engineering Center of Semiconductor Energy-Saving Technology, National Laboratory of Solid-State Microstructures, and Collaborative Innovation Center of Advanced Microstructures. With the support from these laboratories and research centers, the School of ESE has made remarkable achievements. It has undertaken a large number of national scientific and technological research projects, including national 973 Program, 863 Program, and key projects funded by the NSFC. Over the last three years, the average annual research funding exceeds 120 million RMB in total.

In 2016, the School of ESE established its Micro Fabrication and Integration Technology Center as part of the national platform. The facility has four floor space of approximately 5500 m2 cleanroom space. It hosts a large number of first-class processing and testing equipments, compatible with 6-inch wafer process, such as optical and electron beam lithography, metal and thin film deposition, etching, and metrology analysis. Recently, a new femtosecond, extreme ultraviolet, Time Resolved Spin-Angle Resolved Photo-Emission Spectroscopy (TR-Spin-ARPES) beamline was constructed. This 50-fs laser-driven, table-top beamline is integrated with material growth systems, with photon energy ranging from 10 eV to 100 eV to enable surface sensitive studies of the electronic structure dynamics of novel materials.

The School of ESE also establishes international collaboration with top universities, academic institutions and enterprises. Until 2019, many collaborative institutes with government, industry, and universities have been established, some noticeable examples include: NJU-Yangzhou Institute of Optoelectronic Technology, NJU-Zhenjiang Institute of High and New Technology and York-Nanjing Joint Center for Spintronics and Nanoengineering.

Research Highlights

The research program of the School of ESE spans a wide range in electrical engineering and information technology. Here we highlight several research directions.

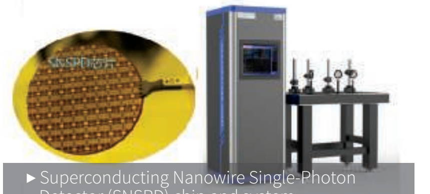
The Key Laboratory of Advanced Photonic & Electronic Materials (LAPEM) is focusing on the research and development of photonic and electronic devices based on wide band-gap semiconductors. The laboratory now has 58 permanent staff, and has built two R&D centers for material growth and characterization platform, and micro-fabrication process lines for Si and compound semiconductor devices. The laboratory members received second prize National Technological Invention Award.

Group of nano-electronics has made several major breakthroughs in two-dimensional electronics, flexible electronics and neuromorphic computing over the past few years. Many papers have been published in Nature Nanotechnology, Nature Electronics and Nature Communications. The group received second prize National Natural Science Award and first prize of Jiangsu Provincial Science and Technology Advancement Award.

Research Institute of Superconductor Electronics (RISE) has been devoting itself for several decades to developing novel superconductor electronic devices and exploring their practical applications. The recent works include several kinds of extremely sensitive detectors from microwave to visible light, quantum computing and quantum microwave circuits. Especially, superconducting nanowire single photon detectors have been used in space communication, space debris detection and biological monitoring. In addition, RISE also develops tunable THz metamaterial based on superconductor, liquid crystal and vanadium dioxide for THz spectroscopic and imaging system, with future applications in wireless communication and biomedicine.

Group of VLSI design works on analog and digital circuits for signal process, 3D NNAD memory and wireless communication. Many breakthroughs have been achieved in the areas of VLSI for High-Speed Communications, Multi-Core Signal Processing, Energy-Efficient AI Processor, and High-Performance Image, Charge, and Acoustic Sensors. Members of the group has been awarded Fellow of IEEE and Best Paper Awards for IEEE VLSI transactions for their outstanding contributions in the field.

Group of Sensing and Imaging Technology focuses on "Big Video System", ranging from spectral video camera, gigapixel imaging, learning-based video/image processing, to gigapixel video streaming. The lab holds a world-class research track record of publications in the top IEEE journals and conferences.



▸ Superconducting Nanowire Single-Photon Detector (SNSPD) chip and system

Finally, the School of ESE warmly welcomes global talents to join us. We offer competitive start-up funding, lab space, salaries and housing subsidy in accordance with the Nanjing University recruiting policy. Outstanding candidates can negotiate case-by-case. For more information please visit the School website: https://ese.nju.edu.cn. Contact: Mr. Jianhong Min (jhmin@nju.edu.cn).

UJS Global Recruiting Program

Jiangsu University is a key comprehensive university founded in August 2001 by merging the former Jiangsu University of Science and Technology, Zhenjiang Medical College and Zhenjiang Teachers' College with the approval of the Chinese Ministry of Education. Jointly supported by Jiangsu Provincial People's Government and the Chinese Ministry of Agriculture and Rural Affairs, Jiangsu University is ranked among the first batch of Jiangsu High-level Universities under Construction, holds the honorary title of National Outstanding Universities in Undergraduate Teaching, and is ranked among the first 50 Sample Universities in China for High Employment of Graduates, one of the first National Sample Universities for Innovation and Entrepreneurship, and one of the first Pilot Universities with Quality Accreditation for International Students Education in China.

Jiangsu University boasts a first-rate teaching level and a batch of high-level disciplines. Its engineering, materials science, clinical medicine, chemistry, agricultural science and pharmacology & toxicology are ranked in the top 1% of the global ESI discipline rankings, coming out 51st in the comprehensive ESI rankings in China's mainland. It has 2 national key disciplines, 1 national key discipline (under cultivation), 10 priority disciplines of Jiangsu colleges and universities, and 7 Jiangsu key first-level disciplines under "the 13th Five-year Plan" and 2 Jiangsu key first-level disciplines (under cultivation) under "the 13th Five-year Plan". The University offers 14 first-level doctoral programs, 44 first-level master programs, and 20 categories of professional master programs, and ME programs in 25 authorized research fields. Moreover, it houses 13 post-doctoral research stations. Jiangsu University enters Top 1000 in internationally reputed rankings such as THE and QS, and ranks among the top 500 in the world, coming out 45th in China's Mainland on Shanghai Ranking's Academic Ranking of World Universities (ARWU). Last but not least, it comes out 41th on the 2019 Comprehensive China University Rankings.

Jiangsu University now employs more than 2,600 full-time teachers, featuring a high-level faculty comprising Distinguished Academicians, as well as experts and scholars sponsored respectively by "Changjiang Scholars Program", and "The National Science Fund for Distinguished Young Scholars". Currently we have more than 37,600 students, comprising more than 23,000 undergraduates, more than 12,000 graduates and nearly 2,000 international degree students.

Forum for High-level Overseas Talents

Forum for High-level Overseas Talents is held every May and December, aimed at creating a platform for the excellent overseas talents to carry out academic exchanges via lectures, seminars and recruitment negotiations, to promote the cooperation of interdisciplinary research, to recruit talents that will help attain the University's objective of building a high-level, research-led university with distinctive features in engineering and internationalization strategy, and to comprehensively advance the University's effort to be ranked among the National "Double First-class" Universities and among Jiangsu High-level Universities.

Recruiting Position

(1) Jinshan Scholars(Type A/Type B)
(2) Jinshan Distinguished Professor(Type A/Type B/Short Term)
(3) Jinshan Youth Distinguished Professor
(4) Competent Professor/Competent Associate Professor
(5) Foreign Postdoctoral Research Fellow

Work and Life Treatment

(1) Competitive remuneration and welfare benefits;
(2) Adequate Start-up Fund and Settling-in Allowance;
(3) Efficient health service provided from campus staff hospital and affiliated hospitals;
(4) Kindergarten and affiliated schools for children of pre-school and compulsory education age;
(5) Furnished interim apartment provided with low-cost rental.

Application Materials

(1) Cover letter for the position that you are applying for;
(2) CV (with publication list);
(3) No less than 3 representative papers;
(4) Expertise and academic results.

How to Apply

All the materials needed for application should be integrated into a PDF document named "name-position-department/school/discipline", to be sent to the email of Human Resources Department (rsk@ujs,edu,cn).

More details for the Jiangsu University Forum for High-level Overseas Talents
Tel: 0086-511-88789658
Email: hr@ujs.edu.cn
Recruitment website: http://rcb.ujs.edu.cn

School Profile

Guizhou Medical University is located in the forest city of Guiyang, the summer capital of the Southwest China. The school was then one of the 9 national medical colleges attached to the Ministry of Education, formerly known as the National Guiyang Medical College, which was founded in 1938. Numerous well-known domestic experts and scholars such as Yang Chongrui, Zhu Zhangyu and Tang Peisong gathered here under the leadership of the first dean, Dr. Li Zongen from Peking Union Medical College and earned it the reputation of"Second Union Medical College in China". The school was renamed Guiyang Medical College in 1950, and renamed Guizhou Medical University in 2015. The school boasts 2 first-level doctoral degree authorization units and 1 professional doctoral degree authorization unit. It is the only provincial key university that has a complete medical talents training system of Bachelor's, Master's and Doctor's degree in Guizhou Province.

The school features in medicine, with coordinate development of science and social sciences (law and education) and literature, engineering and management to support. There are 1 national key laboratory, 2 key laboratories of the Ministry of Education, 2 national experimental teaching demonstration centers and 7 provincial experimental teaching demonstration centers. There are also research platforms of national-level for 2, provincial-level for 25, department-level for 18, and 34 research and innovation teams at all levels. The school has undertaken more than 350 national research projects like the National Natural Science Foundation Program, the National Science and Technology Support Program, the Ministry of Science and Technology Special Fund, the International Science and Technology Cooperation Project for Ministry of Science and Technology, and the National Natural Science Foundation. In 2019, the school made a historic breakthrough for the National Natural Science Foundation declaration work, in which 80 projects got approval, second only to Guizhou University in the province. The clinical medicine discipline has entered the top 1% ranking of ESI worldwide.

We have carried out a comprehensive reform of excellent medical talents training under the full support of the Chinese Academy of Medical Sciences and Peking Union Medical College. The reform is designed to launch several characterized projects like clinical medical innovation class (referred as "Peking Union Class"), small class enrollment, joint training, and full-time tutor system. The school has obvious characteristics and outstanding achievements in the fields of endemic diseases research (Fluorosis disease, Arsenic disease), traditional Chinese medicine ethnic medicine production and research cooperation, pathogenic biology research, stem cell for tissue engineering biomedicine research, and scientific research achievements transformation.

Introduction to the Forum

The first International Youth Scholars Forum of Guizhou Medical University aims to build a platform for exchanges and cooperation among outstanding young talents at home and abroad. It is to promote exchanges and discussions in the field of international leading-edge science and technology research, to advance interdisciplinary and academic innovation. The forum also helps to deepen people's comprehensive understanding of Guiyang and this 80-year-old medical university. Moreover, this forum seeks to bring together and accurately introduce medical talents with lofty ideals who agree with the educational goals, school philosophy, school culture and job demand of Guizhou Medical University.

The forum covers medical-related research fields such as basic medicine, clinical medicine, pharmacy, public health and preventive medicine. This forum is held for the first time, given great concern from the provincial government and relevant departments of Guizhou Province. Excellent young scholars at home and abroad are sincerely invited.

Registration Conditions

1.Excellent overseas talents: under the age of 40; with a doctorate degree from a well-known overseas university.

2. Excellent domestic talents:

a. Under the age of 45; selected into national talent planning, or outstanding young talents with considerable development potential of the same level.

b. With a doctorate degree, under the age of 40, with more than one year of overseas study experience and overseas work experience.

3. If you have a willingness to work full-time, or you want to rely on our school to apply for national talent projects and provincial-level talent projects, you are welcome to cooperate internationally with our school through projects.

Schedule

· Registration: January 2, 2020
· Forum: January 3-4, 2020

1. Opening ceremony and special report meeting;
2. Sub-forum academic seminar;
3. Exchange forum, talent project declaration on-site guidance;
4. Visit School History Museum, related platform labs;
5. On-site contract signing.

Application and Registration

Application accepted from the date of announcement.

Deadline for application: November 30, 2019.

Qualified interested candidates should apply through the "Special Registration Channel for the First International Young Scholars Forum of Guizhou Medical University". scan the QR code below to register in advance:

Activity consultation: Teacher Zhang
Tel: 13810239853 (WeChat) Contact
E-mail: zhangxinyi@eol.cn

Travel and Accommodation

After receiving the formal invitation letter, the invited scholars should order the air ticket or train ticket by themselves. The school will provide the reimbursement according to relevant standards (aircraft: economy class; high-speed train: second-level seat; normal train: hard sleeper). Please keep relevant bills (air ticket invoice, boarding pass, passport copy, ticket, etc.) and hand over to the conference contact person. Room and board during the forum will be arranged by the school.

Contacts: Geng Yan, Li Bingfu, Zhang Xing
Tel: +86-851-88416215
E-mail: fgc@gmc.edu.cn

Hangzhou Innovation Research Institute of Beihang University

Hangzhou Innovation Research Institute, Beihang University is a new high-level research institute jointly established by Beihang University and the Zhejiang Province, Hangzhou City and Binjiang District governments. With the mission of "building a world-class technological innovation platform and innovative talent training platform in the field of information", and focusing on the multidisciplinary intersection of information technology, life and health, cognitive science and new materials, Hangzhou Innovation Research Institute actively explores new mechanisms and gathers global innovative resources, and is committed to achieving a number of major original innovations and key technological breakthroughs and applications, striving to become a talent and innovation center that is rooted into Zhejiang Province while looking to the world's first-class.

In March 2018, Hangzhou Innovation Research Institute of Beihang University officially settled in Binjiang District of Hangzhou, meaning the entering of the second 985 university in Hangzhou. The construction of its graduate school launched in September 2019 at Baima Lake of Binjiang District. The construction of hardware facilities is expected to be completed in about two years, and the scale of graduate students is expected to be 2,000. More importantly, Hangzhou Innovation Research Institute has undertaken the construction work of Sino-French Aviation University, providing teachers reserve for the university. On January 9, 2018, under the testimony of President of PRC Xi Jinping and French President Macron, Beihang University signed the memorandum of cooperation with Ecole Nationale de l'Aviation Civile (the French National Civil Aviation University) and agreed to jointly establish Sino-French Aviation University. The site of the university will be in the town of Pingyao, Yuhang District, Hangzhou, covering an area of 1,500 mu (10,000 acres).

I. Recruitment Positions and Salaries

1. Laboratory Director (with an attractive salary)
2. Qianjiang Excellent Post doctor (annual salary and local comprehensive subsidies totaling 1.07-1.12 million RMB)

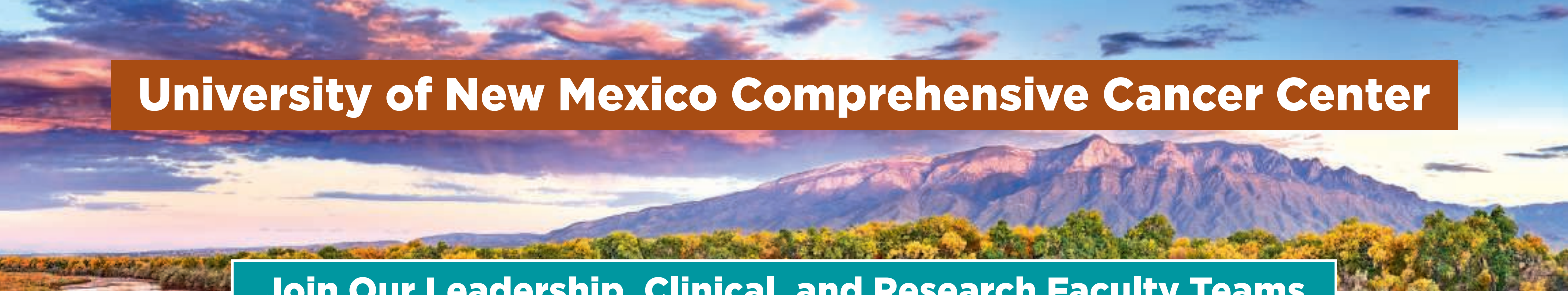
II. Recruitment Majors

Computer Science and Technology, Instrumentation Science and Technology, Optical Engineering, Electronic Science and Technology, Control Science and Engineering, Material Science and Engineering, Information and Communication Engineering, Mechanical Engineering, Physics, Mathematics, Medical Imaging and other related majors.

III. Delivery

This announcement is a long-term recruitment. For details, please refer to the official website of Hangzhou Innovation Institute, Beihang University – recruitment (http://hzii.buaa.edu.cn/zcyz/zpgz.htm). Please send your resume to buaa_hr@buaa.edu.cn (email: "intended position + name")

Contact: Ms. Tian, Mr. Chen
Tel: +86-571-85367559 / 19957890995
E-mail: buaahz_hr@buaa.edu.cn



University of New Mexico Comprehensive Cancer Center

Join Our Leadership, Clinical, and Research Faculty Teams

The University of New Mexico Comprehensive Cancer Center (UNMCCC) is the Official Cancer Center of New Mexico and the only National Cancer Institute (NCI) designated comprehensive cancer center in a 500-mile radius. Our 134 oncology physicians, 122 cancer research scientists, and staff focus on discovering the causes and cures for cancers disproportionately affecting the people of the American Southwest — primarily Hispanic, American Indian, and Non-Hispanic White — with strikingly different patterns of cancer incidence, mortality and disparity. In the past year, our center cared for 12,000 patients; 12 percent participated in therapeutic interventional studies and 35 percent in interventional studies. UNMCCC has outstanding programs in Cancer Control and Population Sciences, Cellular and Molecular Oncology, and Cancer Therapeutics. Our research houses national centers: The Molecular Discovery and High Throughput Target Screening Center (nmmlsc.health.unm.edu), one of six Chemical Biology Consortium Centers of Excellence in The NCI NExT Program; Spatiotemporal Modeling of Cell Signaling (stmc.unm.edu), one of 13 NIH National Centers for Systems Biology; and a NIH Clinical and Translational Sciences Center. We enrich our endeavors by collaborating with Sandia and Los Alamos National Labs and Lovelace Respiratory Research Institute. Benefit from our Shared Resources, which include biospecimen collection and tissue analysis, genomics, biostatistics, bioinformatics, cancer population science and behavioral interventions, and the conduct of clinical interventions. UNMCCC is the center of our statewide cancer clinical trials and health delivery research network — partly funded by a NCI NCORP Grant — and is an Oncology Research Information Exchange Network (ORIENcancer.org) member. Our center has conducted 60+ statewide community-based cancer education, prevention, screening, and behavioral intervention studies involving more than 10,000 New Mexicans. Learn more at cancer.unm.edu.

Associate Director for Cancer Population Sciences

Endowed Chair and Senior Leadership Role

The UNM Comprehensive Cancer Center is searching for a national leader in cancer population sciences. Seeking candidates with a track record of outstanding scholarly achievement reflected in peer-reviewed funding (preferably NCI and NIH), high quality publications, and collaborative interdisciplinary research with a scientific focus in either cancer epidemiology, cancer prevention and control, health services research, behavioral intervention, or cancer health disparities. *Search Chairs: Marianne Berwick and Chuck Wiggins*

Cancer Molecular & Genetic Epidemiology

Endowed Chair and Senior Leadership Role

Seeking cancer population scientists with expertise in population-based molecular and/or genetic epidemiology. Looking for epidemiologists engaged in biomarkers of risk and prognosis, genomics, epigenetics, gene-environment interactions, genetic ancestry, and genetic risk assessment. *Search chairs: Marianne Berwick and Linda Cook*

Cancer Control, Health Services & Behavioral Intervention

Endowed Faculty and Leadership Role

Seeking established population scientists focused on cancer control, health services research, and behavioral intervention research to lead programmatic efforts. Looking for mid-career to senior faculty with outstanding scholarly achievement, including peer-reviewed funding (preferably NCI and NIH) and impactful publications. *Search chairs: Linda Cook and Shiraz Mishra*

Biostatisticians

Seeking PhD biostatisticians to join an outstanding team engaged in statistical methodology relevant to cancer and in biostatistical applications integrated with basic, translational, clinical, and population science research. *Search chairs: Shane Pankratz and Linda Cook*

Cancer Immunology & Tumor Microenvironment

Two Positions: Basic or Translational Scientist

Seeking established mid-career or senior scientists focused on analysis and modeling of pathways that mediate response or resistance to immune therapies, and on signaling perturbations in the context of the tumor microenvironment that enhance or inhibit the immune response to cancer cells. *Search chairs: Eric Prossnitz and Sarah Adams*

Cancer Cell Signaling

Seeking cancer cell biology, signaling, and systems biology experts with interests in dissecting mechanisms of perturbed signaling in cancer cells, on analysis and modeling of pathways mediating response or resistance to targeted therapies. *Search chairs: Diane Lidke and Eric Prossnitz*

RNA Biologist

Seeking highly interactive basic and translational scientists focused on gene expression, transcriptional regulatory and alternative splicing mechanisms relevant to cancer; the biology and role of noncoding RNAs in cancer development and/or progression; and functional genomics (including investigators employing CRISPR/CAS or other functional genomic screening technologies). *Search chair: Scott Ness*

For details and to apply, visit cancer.unm.edu/JoinTheBest

Questions? Contact Search Coordinator Amanda Leigh at ALeigh@salud.unm.edu, (505) 272-2201.

UNM is an Equal Opportunity/Affirmative Action Employer and Educator

Endowed Chairs and Professorships, significant resources, leadership roles, and comprehensive start-up packages available.

Photo: Bill Tondreau, "River Edge", panoramic photographic, courtesy of sumnerdene.com

University of Pittsburgh Tenure-track Faculty Positions in the Department of Structural Biology

The University of Pittsburgh is conducting a broad faculty candidate search for creative individuals who use structural and biophysical methods to address fundamental biomedical questions. The ideal candidate will be motivated to explore applications of his or her structural expertise to disease related questions. We particularly encourage individuals with research activities in cryo-electron tomography for in situ structural biology and solid-state NMR spectroscopy to apply. At present, the cryo-EM facility in the Department comprises 3 Thermo Fisher (FEI) microscopes – a Titan Krios 3Gi equipped with a Falcon 3ec direct electron detector; a TF20 equipped with a TVIPS XF416 camera and Gatan cryoholders; and a T12 equipped with Gatan US 1000 and Orius CCD cameras. The University has committed funding for a Bioquantum/K2 to extend the Krios, and an Aquilos cryoFIB mill, and both will be installed in the spring of 2020. Additional accessory instrumentation is also available. The NMR facility comprises seven Bruker NMR spectrometers ranging from 600-900MHz. Two of them have 89mm bore magnets at 600 and 750Mhz, fully dedicated to state-of-the-art solid-state experiments. The facility is overseen by a full-time manager. The department also possesses dedicated computing resources suitable for processing and storing large datasets.

The University of Pittsburgh is the fifth most highly ranked domestic institution of higher education in terms of NIH funding, and a very wide spectrum of collaborative opportunities exists. The research resources in the Department of Structural Biology and the intellectual environment at the University are truly extraordinary, from state-of-the-art instrumentation to expert support and creative investigators.

Applications at any rank are invited. Successful applicants are expected to develop and lead independent research programs that address important problems in biomolecular systems of wide scientific and medical interest. Competitive salaries and start-up packages will be offered. Applicants should hold PhD and/or MD or equivalent degrees and have demonstrable expertise and scholarly achievement in structural biology or biophysics. The proposed starting date is September 1, 2020 or thereafter. In order to ensure full consideration, applications must be received by **December 31, 2019**.

Application materials including the candidate's curriculum vitae, the names and contact information for three references, and a brief statement of research interests should be emailed to **dxd8@pitt.edu** or sent to:

Dean Duncan
Administrator, Department of Structural Biology
1050 BST3
3501 5th Avenue
Pittsburgh, PA 15260

EEO/AA/M/F/Vets/Disabled.

UNIVERSITY OF SOUTH CAROLINA
School of Medicine

POSTDOCTORAL FELLOWSHIP IN IMMUNOLOGY

Postdoctoral fellowship is available to pursue research supported by NIH grants. Studies will address the epigenomic and microbiomic effects of plant products such as resveratrol, indoles and cannabinoids on inflammation, autoimmunity and cancer. Other projects include studies on the role of dioxins, estrogens and CD44 on immune response. Ph.D. in any biomedical sciences is required with preference given to experience in Immunology.

Apply to Dr. Mitzi Nagarkatti, Chair, Department of Pathology, Microbiology and Immunology, University of South Carolina School of Medicine at this link **http://uscjobs.sc.edu/postings/69253**.

UofSC Columbia is an EOAA Employer and encourages applications from women and minorities.

PURDUE UNIVERSITY

Assistant Professor Position - Biochemistry Department of Chemistry

The Department of Chemistry in the College of Science at Purdue University invites applications for a tenure-track faculty position at the Assistant Professor level in Chemistry - Biochemistry. Successful candidates may have interests in biochemistry as related to human health-relevant biological processes, especially those related to neuroscience or similar fields. If appropriate, the successful candidate will be offered affiliation with Purdue's new Institute for Integrative Neuroscience (PIIN), established as part of the University's strategic investment in the life sciences.

Qualifications: Candidates must have a PhD in biochemistry, or a related field, with outstanding credentials in research, an excellent track record of or potential for leading publications and a strong commitment to excellence in teaching. Successful candidates are expected to develop a vibrant research program supported by extramural funding and to display excellence in teaching at the graduate and undergraduate levels.

The Department and College: With 50 full-time faculty, 350 PhD students, and over 300 outstanding undergraduates, Purdue's highly ranked chemistry department is one of the largest and most diverse in the country. The wide-ranging expertise of the faculty enables fast and effective responses to interdisciplinary research opportunities, positioning it as a key partner in many university-wide centers, institutes, and initiatives. For more information, see http://www.chem.purdue.edu/. Chemistry is part of the College of Science, which comprises the physical, computing and life sciences at Purdue. It is the second-largest college at Purdue with over 350 faculty and more than 6000 students. With multiple commitments of significant investment and strong alignment with Purdue leadership, the College is committed to supporting existing strengths and enhancing the scope and impact of the Department of Chemistry.

Application Procedure: Applicants should submit a cover letter, a curriculum vita, a teaching statement, and a description of proposed research electronically at this site: **https://career8.successfactors.com/sfcareer/jobreqcareer?jobId=8011&company=purdueuniv&username**. Additionally, applicants should arrange for three letters of reference to be e-mailed to the Department Head at biochemsearch@purdue.edu specifically indicating the position for which the applicant is applying. Applications will be held in strict confidence and will be reviewed beginning **December 15, 2019**. Applications will remain in consideration until the position is filled.

A background check will be required for employment in this position. Purdue University's Department of Statistics is committed to advancing diversity in all areas of faculty effort, including scholarship, instruction, and engagement. Candidates should address at least one of these areas in their cover letter, indicating their past experiences, current interests or activities, and/or future goals to promote a climate that values diversity and inclusion.

Purdue University is an EOE/AA Employer. All individuals, including minorities, women, individuals with disabilities, and veterans are encouraged to apply.

Inviting Applications and Nominations for the DEAN OF NATURAL SCIENCES at the University of Puerto Rico, Río Piedras campus

The University of University of Puerto Rico (UPR), Río Piedras, one of the three preeminent research universities designated by the commonwealth of Puerto Rico, is accepting application for the position of Dean of the College of Natural Sciences (Carnegie R2 classification). Candidates should meet the following minimum criteria: Ph.D. in a Natural Science or related discipline; demonstrated competence in research; demonstrated excellence in teaching; proven administrative experience; proficiency in Spanish and English. For additional information and requirements contact (alberto.sabat1@upr.edu).

Please, submit a cover letter articulating your vision for the position; a complete curriculum vitae, and at least three references with full contact information to (**deansearch@upr.edu**).

UPR is an Equal Opportunity, Equal Access Academic Institution that embraces diversity in the workplace.

Ohio State's Pelotonia Institute for Immuno-Oncology Seeks Multiple Tenure-Track Faculty Positions

The Ohio State University Comprehensive Cancer Center – Arthur G. James Cancer Hospital and Richard J. Solove Research Institute (OSUCCC – James) is seeking experienced tenure-track faculty (basic as well as clinical researchers) to work in the recently announced **Pelotonia Institute for Immuno-Oncology (PIIO)** in the following research areas:

- T-cell biology
- Innate immunity and inflammation
- Immune regulation and tolerance
- Cell therapy, synthetic immunology, immunogenomics and cancer vaccines
- Translational and clinical immuno-oncology (IO)
- Immune monitoring and discovery

Led by founding director and renowned immunologist **Zihai Li, MD, PhD**, the PIIO **(cancer.osu.edu/PIIO)** is a comprehensive bench-to-bedside research initiative and represents Ohio State's commitment to grow the exciting area of immuno-oncology.

The **OSUCCC – James (cancer.osu.edu)** is a National Cancer Institute (NCI)-designated comprehensive cancer center and is rated "Exceptional," the highest rating given to cancer centers by the NCI; is one of only a few centers funded by the NCI to conduct phase I and II clinical trials on novel anticancer drugs; and includes the third-largest freestanding cancer hospital in the country.

Candidates will have access to state-of-the-art laboratory spaces, cutting-edge core research facilities, a competitive salary and start-up funding. Applicants must hold advanced degrees, such as an MD, DVM, PharmD, PhD or MD/PhD. The successful candidate is expected to maintain an extramurally funded research program, perform collaborative research, participate in graduate and medical education, be board-certified and meet medical licensure requirements in Ohio if also pursuing a clinical position. Located in Columbus, Ohio **(go.osu.edu/whycolumbus)**, The Ohio State University is an EOE/AA/M/F/D/V employer.

Send cover letter, CV, 2-3 page description of current and future research interests and contact information for three references via email to Tamra Brooks at tamra.brooks@osumc.edu.

Applications will be accepted until the positions are filled.

The James

THE OHIO STATE UNIVERSITY
COMPREHENSIVE CANCER CENTER

Faculty Position
Yale Stem Cell Center
Yale University School of Medicine

LUX ET VERITAS

The Yale Stem Cell Center invites applications for a faculty position at the rank of Assistant or Associate Professor. Rank and tenure will be commensurate with experience. Applicants should have a Ph.D. and/or M.D. degree. The successful candidate will be expected to develop a vigorous, externally funded research program on fundamental questions related to the biology of embryonic or adult stem cells. The investigator will establish a lab in a state-of-the-art new building and will join a vibrant stem cell research community at Yale with over 100 labs working on various aspects of stem cell biology and regenerative medicine. The investigator will also contribute to teaching graduate and/or medical students as well as shaping stem cell research at Yale.

Applicants must apply through the Interfolio website at: **http://apply.interfolio.com/70014** by **December 30, 2019**. Inquiries about the position may be sent to Dr. Haifan Lin, Director of the Yale Stem Cell Center, c/o Madeline Riccio (madeline.riccio@yale.edu).

Yale University is an Affirmative Action/Equal Opportunity Employer. Yale values diversity among its students, staff, and faculty and strongly welcomes applications from women, persons with disabilities, protected veterans, and underrepresented minorities.

CALL FOR PAPERS
Contributed Paper Submission Deadline:
January 23, 2020

The IEEE Engineering in Medicine and Biology Society is pleased to announce the 42nd International Engineering in Medicine and Biology Conference, to be held in Montreal, Quebec, Canada, July 20th - 24th, 2020.

http://embc.embs.org/2020/

WORKING LIFE

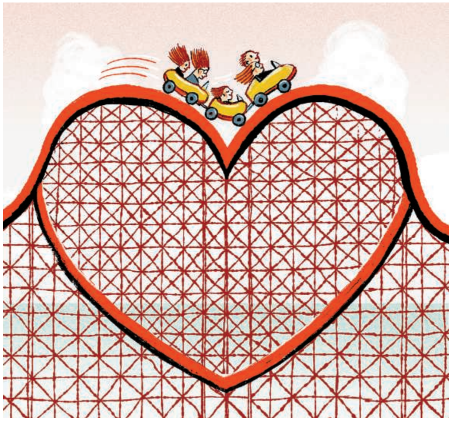
By **J. Andrew DeWoody** *and* **Yssa D. DeWoody**

Our unexpected ride

Life was on track just as we had planned: two university jobs, with tenure for Andrew and a clear path toward a tenure-track position for Yssa (check); the purchase of our first home (check); and three beautiful daughters (check). But 3 months after our youngest daughter Marie was born, she started to have seizures. After a terrifying ambulance ride and a nerve-wracking week in a children's hospital, Marie was diagnosed with ring chromosome 14 syndrome, an exceedingly rare and debilitating condition caused when the 14th chromosome is fused into a ring. Life for our family was about to switch tracks, and we were along for the ride but no longer in control.

Marie's seizures had no regard for class schedules or family time. We tried to keep up with our teaching duties, but it was difficult to lecture knowing that our young daughter could have a life-threatening seizure at any moment. Neither of us wanted to give up our academic career—and we didn't want the other to have to, either. But rigid teaching schedules proved untenable, especially when coupled with the other stresses of academic life, and we needed more flexibility.

The reality was that Andrew was tenured, with the associated job security, and Yssa had applied mathematics expertise that made her more marketable outside academia. After discussing our options, we decided that Yssa would forgo her academic appointment. She obtained a part-time research position in industry, which required fewer hours for equivalent pay and allowed her to work from home when needed. She also started a nonprofit organization for families and researchers interested in ring 14. Our lives were still complicated—including multiple therapy sessions each week plus specialized, expensive child care—but it seemed we had found a way forward.



"When Marie was born our smooth, straight ride ... turned into a roller coaster."

We made this arrangement work for many years. But as our older daughters approached college age, we needed to re-evaluate our situation again. They had matured into fantastic caregivers, and their departure would have a huge impact on their sister and on us. Marie's seizures continued unabated, she was hospitalized regularly, and new concerns were looming as she approached adolescence. Facing numerous surgeries and complex long-term care, we struggled to achieve the elusive work-life balance and needed more support than well-meaning friends could offer.

So, 5 years ago, we moved from Indiana to our native Texas, where most of our extended family resides. We developed a plan to sell the move to Andrew's department head and dean: He'd take a 1-year sabbatical at 50% pay while spending the other half of his time working in an unrelated family business. That way we could give an in absentia appointment a try.

The administrators were understandably skeptical at first, but they agreed to a trial run. The first year of Andrew's "sabbatical" ultimately proved successful, so he resigned from the family business and returned full-time to his faculty role—buoyed by the support of empathetic students, colleagues, and administrators. He spends about 1 week per month on campus and continues to teach, maintain a research lab, and serve on myriad committees. It's often difficult, as Andrew sometimes struggles to arrange online access to meetings and he misses having colleagues down the hall. But so far it has been tenable.

By helping us find a better work-life balance, the arrangement has also allowed Yssa to continue her work, which is now fully focused on running national and international ring 14 nonprofits. Yssa misses many aspects of her prior professional life, most of all the students. However, she finds intellectual stimulation and reward in working with parents and physicians to push medicine forward and help families become their own advocates.

When Marie was born our smooth, straight ride unexpectedly turned into a roller coaster, with many joys but also with some poignant regrets. Ultimately, though, we are fortunate in many ways: wonderful children, a strong marriage, and rewarding work aided by a university that accommodates our whole family's needs. ■

J. Andrew DeWoody is a professor at Purdue University in West Lafayette, Indiana. Yssa D. DeWoody is president and CEO of Ring14 USA and Ring14 International.

ILLUSTRATION: ROBERT NEUBECKER

Published by AAAS